



**PHD**

**The role of scopoletin in cassava post-harvest physiological deterioration**

Liu, Shi

*Award date:*  
2017

*Awarding institution:*  
University of Bath

[Link to publication](#)

**Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

**Take down policy**

If you consider content within Bath's Research Portal to be in breach of UK law, please contact: [openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk) with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.

# **The role of scopoletin in cassava post-harvest physiological deterioration**

Shi Liu

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Biology and Biochemistry

September 2016

## **COPYRIGHT**

Attention is drawn to the fact that copyright of this thesis rests with its author. A copy of this thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and they must not copy it or use material from it except as permitted by law or with the consent of the author.

This thesis may be made available for consultation within  
the University Library and may be photocopied or lent to other libraries  
for the purposes of consultation.

Signed:

Shi Liu

# CONTENTS

Contents .....	2
Acknowledgements .....	7
Declaration of work done in conjunction with others .....	9
Abstract .....	10
Abbreviations .....	11
1 Introduction .....	15
1.1 Cassava, a tropical crop .....	15
1.1.1 What makes it a cassava .....	15
1.1.1.1 Background .....	15
1.1.1.2 Physiology .....	16
1.1.1.3 Propagation and cultivation .....	20
1.1.2 The history of cassava cultivation .....	21
1.1.3 Importance of cassava in human life .....	23
1.1.4 Why has cassava not been used wider? .....	24
1.1.4.1 Cyanogenic compounds .....	24
1.1.4.2 Low nutritional value of cassava .....	25
1.1.4.3 Diseases .....	26
1.2 Post-harvest physiological deterioration (PPD) .....	27
1.2.1 Post-harvest physiological deterioration and its importance .....	27
1.2.2 Attempts to alleviate PPD .....	29
1.2.2.1 Delaying the harvest .....	29
1.2.2.2 Pruning .....	30
1.2.2.3 Exclusion of oxygen .....	30
1.2.2.4 Curing .....	31
1.2.3 The mechanism of PPD .....	31
1.2.4 Scopoletin and phenolic secondary metabolites in PPD .....	35
1.3 Scopoletin synthesis and phenylpropanoid metabolism .....	37
1.3.1 Phenylpropanoid metabolism in higher plants .....	37
1.3.2 Scopoletin biosynthesis involves three pathways in phenylpropanoid metabolism .....	39

1.4 Stress responses in other root and tuberous crops.....	42
1.4.1 Potato ( <i>Solanum tuberosum</i> ).....	42
1.4.2 Sweet potato ( <i>Ipomoea batatas</i> ).....	42
1.4.3 Yam ( <i>Dioscorea</i> spp.).....	43
1.4.4 Taro ( <i>Colocasia esculenta</i> ) .....	44
1.4.5 What can we learn? .....	44
1.5 Research strategy.....	45
2 Instruments, materials, and methods.....	46
2.1 Instruments and chemicals .....	46
2.2 Plant materials .....	46
2.3 Bacteria strains.....	47
2.4 Plasmids.....	47
2.5 Software and statistical methods .....	48
2.6 Primers.....	49
2.7 Bacteria and plants manipulation .....	49
2.7.1 Antibiotics .....	49
2.7.2 Culture of bacteria .....	50
2.7.3 Preparation of electro-competent <i>A. tumefaciens</i> .....	50
2.7.4 Transformation of chemically competent <i>E. coli</i> .....	51
2.7.5 Transformation of electro-competent <i>A. tumefaciens</i> .....	51
2.7.6 Growth of <i>Arabidopsis</i> .....	51
2.7.7 Homozygous T-DNA inserted <i>Arabidopsis</i> selection.....	52
2.7.8 <i>Arabidopsis</i> transformation and homozygosity selection .....	53
2.7.9 <i>Arabidopsis</i> Gus ( $\beta$ -glucuronidase) staining.....	54
2.7.10 Subculture and growth of cassava .....	54
2.7.11 Transformation of cassava .....	55
2.8 Nucleic acid manipulation.....	55
2.8.1 <i>Arabidopsis</i> leaf genomic DNA extraction .....	55
2.8.2 <i>Arabidopsis</i> RNA extraction .....	56
2.8.3 Cassava leaf DNA extraction .....	56
2.8.4 Southern blotting .....	57

2.8.5	Cassava root RNA extraction.....	57
2.8.6	<i>E.coli</i> plasmid extraction .....	58
2.8.7	cDNA preparation and quality control.....	59
2.8.8	PCR .....	59
2.8.8.1	Genotyping/screening PCR.....	59
2.8.8.2	Colony PCR .....	60
2.8.8.3	Cloning PCR products.....	61
2.8.8.4	'Touch-up' PCR.....	61
2.9	Plasmid manipulation.....	62
2.9.1	DNA restriction endonuclease digestion .....	62
2.9.2	DNA electrophoresis and clean-up .....	62
2.9.3	A-tailing and TOPO <sup>®</sup> TA cloning.....	63
2.9.4	DNA ligation .....	64
2.9.5	Gateway <sup>®</sup> cloning .....	64
2.9.6	DNA sequencing .....	66
2.10	PPD discolouration quantification .....	66
2.10.1	Cassava root sample preparation .....	66
2.10.2	Quantification of discolouration .....	67
2.11	LC-MS biochemical quantification .....	68
2.11.1	<i>Arabidopsis</i> root sample preparation .....	68
2.11.2	Cassava root sample preparation .....	68
2.11.3	Scopoletin extraction and LC-MS quantification .....	68
2.12	Quantitative RT-PCR and data analysis.....	69
2.12.1	Primer efficiency quality control.....	69
2.12.2	Quantitative PCR .....	70
3	Do F6'H1 candidate genes exist in cassava genome?.....	71
3.1	Introduction .....	71
3.1.1	Scopoletin and PPD .....	71
3.1.2	Scopoletin biosynthesis in cassava and <i>Arabidopsis</i> .....	72
3.1.3	<i>Arabidopsis</i> scopoletin biosynthesis and F6'H1.....	73
3.1.4	Aim of research .....	74

3.2 Results.....	75
3.2.1 Cassava has a small family of F6'H1 candidate genes .....	75
3.2.2 At least four of the candidate genes were expressed in cassava roots during PPD.....	78
3.2.3 The phylogenetic tree revealed the relation between the cassava candidates gene family and F6'H1 .....	82
3.2.4 The genes to be used in downstream experiments were selected .....	84
3.3 Discussion.....	88
4 Do cassava F6'H1 candidates complement scopoletin biosynthesis in <i>Arabidopsis</i> F6'H1 T-DNA insertion mutants?.....	89
4.1 Introduction .....	89
4.1.1 Background .....	89
4.1.2 Strategy .....	89
4.2 Results.....	90
4.2.1 Confirmation of the <i>Arabidopsis</i> F6'H1 T-DNA insert mutants .....	90
4.2.2 Cloning cassava cDNA of F6'H1 candidates into T-DNA vectors driven by suitable promoters.....	91
4.2.3 <i>Arabidopsis</i> T-DNA insert lines were transformed with cassava F6'H1 candidate genes .....	98
4.2.4 Determination of scopoletin levels in wild-type, T-DNA insert mutants <i>Arabidopsis</i> lines .....	100
4.2.5 Confirmation of the functional identity of cassava candidate genes .....	102
4.3 Discussion.....	106
5 Does inhibiting 6'-hydroxylation of feruloyl CoA in cassava roots alter the biosynthesis of scopoletin and PPD response?.....	108
5.1 Introduction .....	108
5.1.1 Post-harvest physiological deterioration (PPD) .....	108
5.1.2 RNAi inactivated cassava gene expression .....	109
5.2 Aims .....	110
5.3 Results.....	110
5.3.1 Cassava candidate gene 10376 was used to construct RNAi.....	110
5.3.2 Wild-type cassava was transformed .....	111
5.3.3 RNAi altered cassava phenotypes and PPD discoloration .....	114

5.3.4	RNAi down-regulated the expression of the cassava F6'H1 candidates ..	119
5.3.5	RNAi down-regulated scopoletin accumulation .....	122
5.3.6	PPD development and scopoletin accumulation are related to the expression of cassava F6'H1 gene family .....	124
5.4	Discussion.....	126
6	General discussion .....	129
6.1	The role of scopoletin on a broader horizon .....	129
6.2	The cassava F6'H1 candidates are highly PPD-inducible .....	132
6.3	Source and sink of scopoletin .....	132
6.4	Future work.....	134
7	References.....	137
8	Appendix .....	151
	Appendix I Primers and their details .....	151
	Appendix II The recipe of solutions used in this project .....	158
	1, LB medium.....	158
	2, 2YT medium .....	158
	3, SOC medium.....	159
	Appendix III Enlarged phylogenetic tree of <i>Arabidopsis</i> and cassava F6'H1 candidate genes .....	160
	Appendix IV Full length amino acid alignment of the seven cassava candidates genes.....	162

# ACKNOWLEDGEMENTS

During the past four years of research many people have helped me. Without their help it would have been impossible for me to reach the end of my study and start thinking about my next step in life.

My greatest thanks are to my parents who funded my studying during this project. Although they could not be here with me all the time, their encouragement is with me every day. I thank the Department of Biology and Biochemistry, University of Bath for providing a tuition fee waiver to me. I was also helped by the Biochemical Society and the World Congress on Root & Tuber Crops, who made it possible for me to attend the Third Scientific Conference of the Global Cassava Partnership for the 21st Century & 17th Symposium of the International Society for Tropical Root Crops, and discuss my work with colleagues from the entire world.

Dr John R. Beeching and Dr James Doughty, my two supervisors, should have my gratefulness. They are the ones guiding me through the four years of study. Where there were difficulties, there were their encouragement and helpful advice. During the write-up stage I especially benefitted from the tireless efforts that Dr Beeching had spent on helping me write this thesis.

The transgenic cassava and the PPD scoring software used in this project were developed by our co-operators at ETH-Zurich, Dr Herve Vanderschuren, Dr Norbert Kirchgessner and Dr Ima M. Zainuddin; they not only gave me highly valued help, but also contributed considerably to this project.

Dr Andrew Preston and Dr Jerry King helped me construct the CaMV 35S Gateway<sup>®</sup>-competent plasmid for my project, as well as kindly letting me use their MicroPulse electroporator and also their StepOnePlus<sup>™</sup> quantitative PCR machine; thank Professor Michael Danson for allowing me to use his NanoVue spectrophotometer for my endless RNA samples. Dr Simon E. Bull constructed the StPAT Gateway<sup>®</sup>-competent plasmid long time ago, and scavenging his legacy in the -80°C freezer made my life easier. I should also thank Dr Mike T. Page and Dr Kim Reilly, whose previous work inspired me in many ways.

Thanks to Dr Nor H. Mahmod for giving me my introduction to the lab as well as enabling me to start my lab work. Whenever I had difficulties with my research, her



suggestions were always available; to Lidia K. Al-halaseh and Yi-wen 'Fern' Lim whose work on LC-MS and qRT-PCR benefited me; to Thoni, Alex, Aswad, Trupti, Glenda, and everyone in 4 South 1.52, thank you all for your academic and non-academic support.

Dr Anneke Lubben operated the LC-MS scopoletin quantification for us. I did not realise how much work she did until I had the chance operating it myself.

I also need to say thank you to Dr Maha Al-Jabri for training me in the basic skills of handling *Arabidopsis*, and to Dr Julia Tratt for providing the wild type Columbia-0 *Arabidopsis* seeds. Dr Baoxiu Qi provided *Agrobacterium tumefaciens* and *E. coli* DH5 $\alpha$  strains for this project. Yaxiao Li taught me how to perform the *Arabidopsis* floral dipping. Besides, lab 3S 1.13 is like my second base, and whenever there was a problem could not be resolved 'locally', they were my first choice to consult.

## **DECLARATION OF WORK DONE IN CONJUNCTION WITH OTHERS**

Dr Simon E. Bull constructed the Gateway<sup>®</sup>-competent vector with StPAT promoter; Dr Andrew Preston and Dr Jerry King helped construct the Gateway<sup>®</sup>-competent vector with CaMV 35S promoter; Dr Ima M. Zainuddin transformed the RNAi vectors into wild-type cassava TMS 60444, and verified the single-inserted RNAi transgenic cassava lines; Dr Norbert Kirchgessner and Dr Herve Vanderschuren designed the cassava PPD discoloration quantification software.

# ABSTRACT

Cassava (*Manihot esculenta* Crantz) is an important tropical crop which provides a large portion of daily calories intake to hundreds of millions of people in Africa, Latin America, and tropical Asia. Cassava is grown for its starchy storage roots as staple food, as animal feed, and as industrial raw material. The utilisation of cassava is hindered by its characteristic physiological response, the post-harvest physiological deterioration (PPD). The inevitable wounding caused during harvesting and handling will trigger a series of physiological responses within 24 to 48 hours, which causes a blue-black discoloration in the storage roots, rendering these roots unmarketable and unpalatable in a few days. During the PPD response large amount of phenylpropanoid compounds, especially scopoletin and its glycoside, accumulate in the roots. Scopoletin may play an important role in PPD development but little work has been done on the possible relationship. Here we aim to examine the effects of altering scopoletin synthesis in cassava roots on the PPD response.

In *Arabidopsis thaliana*, gene F6'H1 (feruloyl CoA 6'-hydroxylase 1) is indispensable in the biosynthesis of scopoletin. Cassava F6'H1 candidate gene family involved in scopoletin synthesis were identified by their ability to functionally complement F6'H1 T-DNA insertion mutation in *Arabidopsis thaliana* that prevented synthesis of scopoletin. RNAi constructs targeting the identified cassava F6'H1 candidate gene family were designed, under the control of either constitutive CaMV 35S or root-specific *StPAT* promoters. These were used to transform wild-type cassava to down-regulate the expression of these scopoletin synthetic genes in F6'H1 gene family. The inhibition of cassava F6'H1 candidate gene expression and thus the scopoletin synthesis in transgenic cassava roots were confirmed by qRT-PCR and LC-MS, respectively. The RNAi transgenic cassava lines show less scopoletin accumulation and inhibited F6'H1 candidate genes expression during the PPD response. A reduced PPD discoloration development compared to that of the wild-type was also observed in the RNAi transgenic cassava lines.

# ABBREVIATIONS

°C	degree Celsius
ABA	Absciscic acid
<i>Agrobacterium</i>	<i>Agrobacterium tumefaciens</i>
ANOVA	analysis of variance
AOX	alternative oxidase
APX	ascorbate peroxidase
<i>Arabidopsis</i>	<i>Arabidopsis thaliana</i>
Att	attachment site
bp	base pair
CaMV	cauliflower mosaic virus
CAT	catalase
CAM	cassava axillary medium
CAMBIA	Centre for the Application of Molecular Biology to International Agriculture
CBM	cassava basal medium
CCoAOMT	caffeoyl CoA O-methyltransferase
<i>C. elegans</i>	<i>Caenorhabditis elegans</i>
C4'H	cinnamate 4'-hydroxylase
CIAT	Centro Internacional de Agricultura Tropical
CTAB	cetrimonium bromide
C3'H	caffeate 3'-hydroxylase
Cu	cupric
DAH	days after harvest

DEPC	diethyl pyrocarbonate
DMSO	Dimethyl sulfoxide
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylene diamine tetra acetic acid
ETH	Eidgenössische Technische Hochschule (Swiss Federal Institute of Technology)
EtOH	ethanol
Fe	iron
F6'H	Feruloyl CoA 6'-hydroxylase
FW	Fresh weight
g	gram
GM	genetically modified
GUS	β-glucuronidase
hptII	hygromycin phosphotransferase II
iTRAQ	isobaric tags for relative and absolute quantification
LB	lysogeny broth
LC-MS	liquid chromatography-mass spectrometry
M	molar concentration
MeOH	methanol
MES	2-(N-morpholino) ethanesulfonic acid
mg	milligram
Mins	Minutes
ml	millilitre
MS medium	Murashige-Skoog medium

NASC	Nottingham <i>Arabidopsis</i> Stock Centre
NCBI	The National Center for Biotechnology Information databases
nt	nucleotides
PAL	phenylalanine ammonia-lyase
PCR	polymerase chain reaction
Phe	phenylalanine
PVP	polyvinylpyrrolidone
POX	peroxidase
PPD	post-harvest physiological deterioration
ppm	part per million
qRT-PCR	quantitative reverse transcription PCR
RNAi	RNA interference
ROS	reactive oxygen species
RQ	relative quantity
RT	reverse transcription
SDS	sodium dodecyl sulfate
Secs	seconds
SOC	Super Optimal broth with Catabolite repression
SOD	superoxide dismutase
StPAT	<i>Solanum tuberosum</i> PATATIN
TAE	tris-acetate EDTA
TAN	tropical ataxic neuropathy
TMS	tropical <i>Manihot</i> series
Tris	tris (hydroxymethyl) methylamine

UV	ultraviolet
V	volt
WHO	world health organization
WT	wild-type
X-Gluc	5-bromo-4-chloro-3-indolyl glucuronide
2OGD	2-oxoglutarate-dependent dioxygenase
4CL	4-coumarate coenzyme A ligase
4-MU	4-methylumbelliferon

# 1 INTRODUCTION

## 1.1 CASSAVA, A TROPICAL CROP

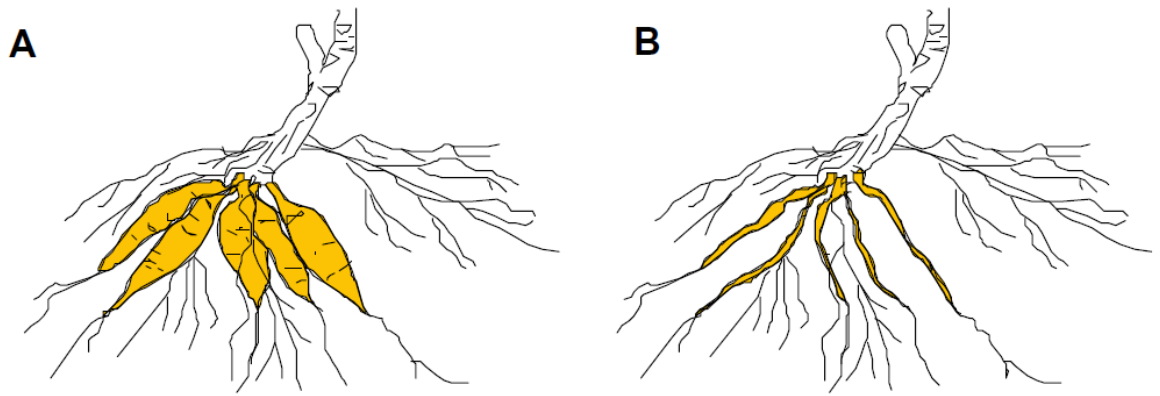
### 1.1.1 What makes it a cassava

#### 1.1.1.1 Background

Cassava (*Manihot esculenta* Crantz.) is a tropical crop in the family Euphorbiaceae (Blagbrough et al., 2010). Cassava, whose English common name probably came from *casabe*, an indigenous South American flatbread made from cassava starch (Benito et al., 2013, Clement et al., 2010), has many different names in diverse cultures, ranging from manioc in French-speaking Africa, to muhogo in Swahili-speaking Africa (Bull et al., 2011), to yuca in Mexico, to mandioca in Brazil, and to mushu (literally *tree-potato*) in Mandarin Chinese.

In the genus *Manihot* there are 98 species and many of them had been utilised by humans as a source of food (Olsen and Schaal, 2001). In the past it was believed that the species *M. esculenta* was generated by interbreeding between different species in the genus *Manihot* (Allem, 1994), but a study on the microsatellite polymorphism of different *Manihot* species showed that the microsatellites of domestic *M. esculenta* are a subset of that of a related wild subspecies, *M. esculenta flabellifolia*, indicating that *M. esculenta* was domesticated from the closely related wild subspecies solely (Olsen and Schaal, 2001). During the domestication process *M. esculenta* acquired enlarged storage roots compared to that of the wild-type, as shown in figure 1.1.





**Fig. 1.1** The roots of a domesticated *Manihot esculenta* (A) and a wild-type *M. esculenta flabellifolia* (B). The storage roots are highlighted. Picture modified from (Jansson et al., 2009).

#### 1.1.1.2 Physiology

Cassava ( $2n=36$ ) is a perennial plant growing into a shrub or a small tree depending on branching habit (Blagbrough et al., 2010, El-Sharkawy, 2004), with a fully grown size of about 1-4 meters in height (Alves, 2002). A cassava plant is shown in figure 1.2. During maturation the stem will gradually lignify from the base, while the top part is still herbaceous and green.

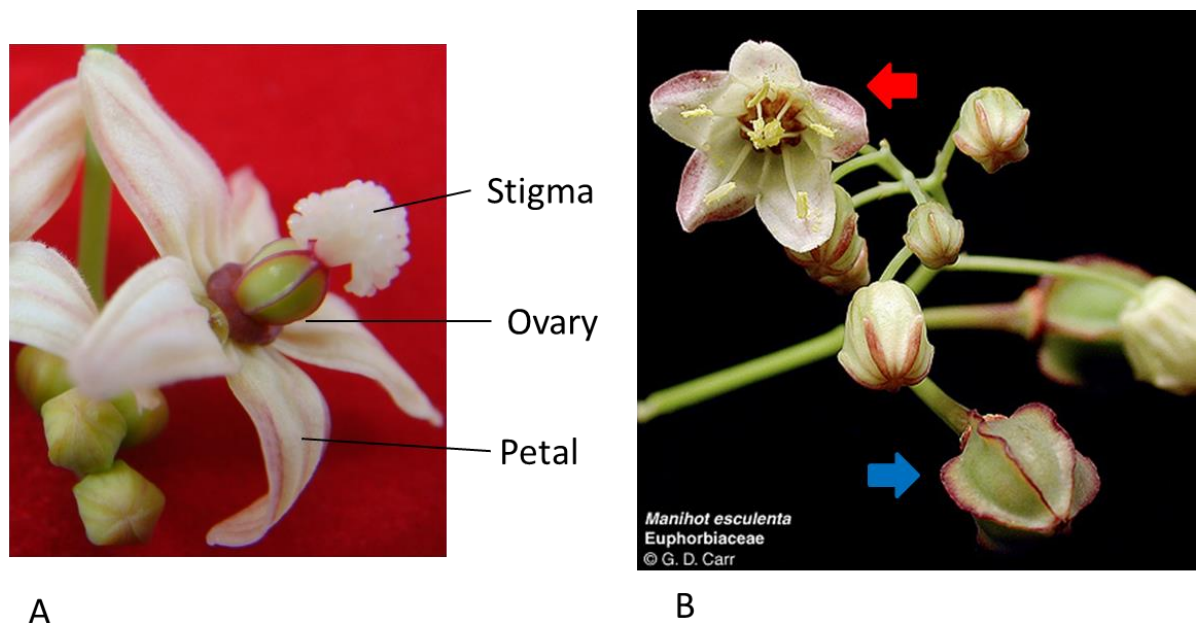
Like many Euphorbiaceae plants, cassava is monoecious, bearing flowers of both genders on the same inflorescence, that emerges from the joint of reproductive branches (Alves, 2002), as shown in figure 1.3. On the inflorescence male flowers form on the top and female flowers are on lower parts. During flowering, female flowers open 1-2 weeks before the males, encouraging cross pollination (by insects). Due to this, cassava population is highly heterozygous. Whether or not a cassava plant will flower is unpredictable, and the time of flowering is often staggered among individuals (Ceballos et al., 2004). In some cultivated varieties flowering has never been recorded (Manu-Aduening et al., 2005).



**Fig. 1.2 A Cassava plants growing in the field.** Photo courtesy of Simon E. Bull.

Cassava grows simple palmate leaves with palmate veins, and each leaf develops three to nine lobes depending on the age and growth condition of the leaves (Aloys and Ming, 2006). The leaves are in alternate arrangement, attached to long and slender petioles (approximately 5-30 cm, depending on the size of the plant). The upper surface of a cassava leaf is protected by a layer of waxy epidermis. The stomata are mainly located on the lower surface, implying a hot place of origin (Alves, 2002).

A successfully fertilised ovary will develop into a globular-shaped capsule (figure 1.3 B) containing three seeds with a diameter about 1-1.5 cm within 90 days. The seeds have a shape similar to castor beans, with brown dotted pattern on the smooth seed coat (Alves, 2002). The seeds are dormant for a few months after maturity and their germination requires a high temperature, rendering propagation by seeds difficult (Ceballos et al., 2004).

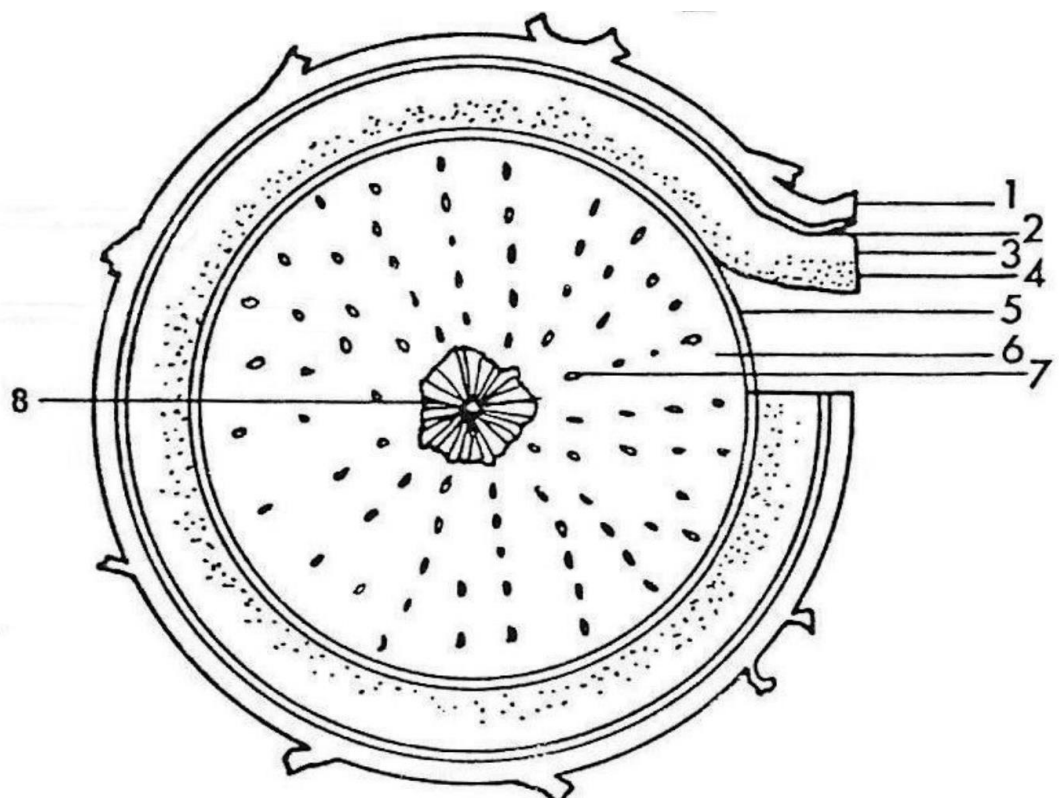


**Fig 1.3 Flowers and a blooming inflorescence of cassava.** (A) A fully opened female flower (right) and unopened male flowers (background). (B) Part of an inflorescence. Male flowers are on the top (fully opened, red left arrowhead) and female flowers are lower (fertilised, blue right arrowhead). Photos courtesy of Ahmad Fathoni and Gerald D. Carr.

During the growth of a cassava plant some fibrous roots will develop into storage roots. Depending on the cultivation variety and the growth condition, 3-10 tuberous roots will form, although occasionally more than 10 storage roots will develop on one individual. Generally an individual plant yields 2-3 kg of storage roots (van Oirschot et al., 2000). A mature storage root is conical to carrot-shaped, with a length up to 100 cm and a diameter from 3 to 15 cm (Alves, 2002). The developed storage root is comprised of three different tissues: a brown periderm ('bark'), a white and thin cortex ('peel'), and the parenchyma (the edible part) with central and scattered vascular bundles in the flesh. The structure of a cassava root is shown in figure 1.4. The periderm is comprised of a thin layer of dead cork cells. During the thickening of the root in soil, the periderm keeps tearing and wearing off, being replaced by newly formed cork cells beneath it. The cortex is flexible and about 1 mm thick, containing sclerenchyma, cortical parenchyma and phloem (Alves, 2002). The cortex does not adhere firmly to the parenchyma, and it can easily be peeled off from the flesh. The parenchyma, which accounts for about 85% of the root's weight, is an accumulation of parenchymatous cells containing a large amount of

starch. Small bundles of vascular ducts are present throughout the parenchyma, and a large vascular bundle runs through the tuberous root at the centre (Onwueme, 1978). The roots develop from fibrous roots via secondary thickening, which is possibly a trait acquired through domestication (Jansson et al., 2009). During the initiation of storage roots, activity of the cambium of the 'fated' fibrous roots will be up-regulated rapidly, differentiating the secondary xylem into enlarged parenchyma cells. Starch will accumulate in these cells (Zhang et al., 2003b). Thus, a storage root develops purely from root tissue, and is not able to sprout or reproduce. In other words, cassava roots are not propagules.

The formation of storage roots requires short-day light. With daylight time longer than 10-12 hours per day, storage root formation will be inhibited. Due to this reason, the preferred region for cassava cultivation is the low latitude tropics (Onwueme, 1978). Under optimal conditions, the enlargement of storage roots begins as early as eight weeks after planting.



**Fig. 1.4 A transverse section of a cassava storage root.** (1) periderm, (2-4) the peel, (2) sclerenchyma, (3) cortical parenchyma, (4) phloem, (5) cambium, (6) storage parenchyma, (7) scattered vascular bundles, (8) central vascular bundle. Picture modified from (Hunt et al., 1977)

### 1.1.1.3 Propagation and cultivation

In field cultivation, vegetative propagation is usually used. Propagation by seeds is possible, but the germination is difficult and the growth of seedlings is slow, therefore is not preferable (Alves, 2002, Onwueme, 1978). In addition, due to cross-pollination, traits of the parents will segregate in seedlings.

Unlike many other root and tuberous crops, such as potato and yam, cassava roots cannot be used to propagate the plant, because technically it is a true root that does not have buds. Therefore, stem cuttings are used as propagating material (Alves, 2002). After harvesting, the woodened section of cassava stem is cut into pieces which provide the axillary buds as well as organic and inorganic nutrients to initiate the growth of plantlets. To obtain the best propagating quality, hardened but not yet extensively lignified stems from cassava of 8-18 months old are used (Lozano, 1977). The cuttings should be 20-30 cm in length, possessing 5-7 nodes to provide enough nutrients and budding viability. Using longer cuttings (50-60 cm) may increase the yield but also make planting and harvesting difficult (Onwueme, 1978). The ideal conditions to store the stem cuttings are 20-23 °C, 80% relative humidity, and with a fungicidal treatment before storage (Lozano et al., 1977). With proper storage conditions, the viability of stem cuttings can last for one month before planting.

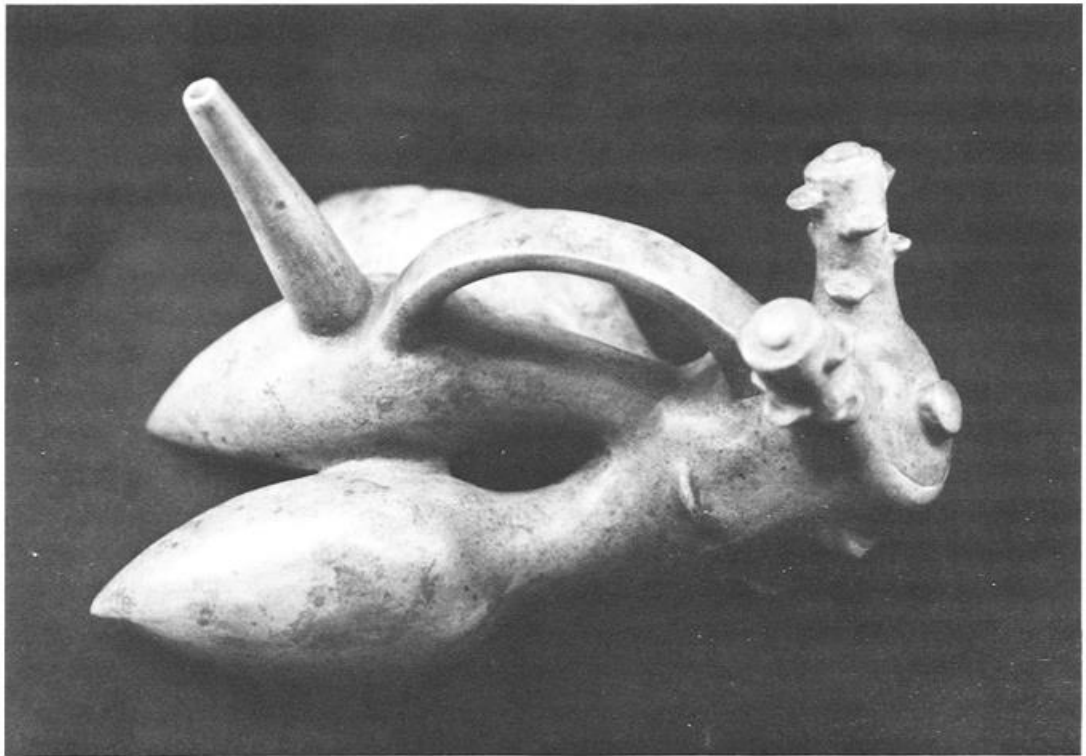
The stem cuttings need to be planted in soil with adequate moisture. Five to seven days after planting, adventitious roots will emerge from nodes of axillary buds, while the bottom of a cutting will start forming a callus, from which rootlets will develop (El-Sharkawy, 2004). However, if the seedling is grown from a seed, a tap root system will form. Ten to thirteen days after planting, first leaves will have developed (Jansson et al., 2009). During the first two months of a cassava's life, it produces a well-developed system of roots, stems and leaves, with little visible storage root formation. From the third month after planting, about 3-10 fibrous roots will start thickening, developing as storage roots (Alves, 2002), as the stems and leaves grow rapidly. By 120-150 days after planting, the plant will achieve its maximal canopy, starting to produce far more photosynthetic products than it needs. The carbohydrate is stored in the storage roots as starch. As a result, from the sixth to seventh month after planting, the development of storage roots will accelerate, driven by the increasing amount of carbohydrate transported from leaves to the roots (Alves, 2002).

As a tropical crop, cassava grows best in a warm and humid climate, with an optimal average temperature around 25-29 °C (Onwueme, 1978). If the temperature falls below 17°C or rises above 37°C, its growth will halt (El-Sharkawy, 2004). Annual rainfall of 1000 to 1500 mm is ideal for cassava growing, but the plant is tolerant to drought and able to grow and yield when water supply is limited. The best soil for cassava cultivation is lightly sandy loam with good drainage. Although cassava grows better in highly fertile soil, it also produces an acceptable yield when the fertility is low and the soil is acidic (El-Sharkawy, 2004). Due to its high yield and low input requirements, cassava is adopted by many poor and small-holder farmers as a main source of food, which gives cassava a stigmatised title of 'the third world food'.

### **1.1.2 The history of cassava cultivation**

Evidence suggests that the domestication of cassava began at approximately seven thousand years ago in the Amazon rainforest region, South America (Peroni et al., 2007). In the past there were different opinions, arguing cassava might have been domesticated either in Amazon basin or in Mesoamerica (Sambatti et al., 2001). A microsatellite polymorphism study (Olsen and Schaal, 2001) supported the hypothesis that cassava was domesticated only once in a single region instead of in parallel. A phylogeographic study on cassava G3pdh (Glyceraldehyde 3-phosphate dehydrogenase) gene haplotypes suggested that the wild population of cassava in the southern Amazon basin shared the most significant similarity with the domesticated cassava (Olsen and Schaal, 1999).

A study on starch grains remaining on pre-historical stone tools from Panama showed cultivation of cassava had been introduced to Mesoamerica 5000-6000 years ago (Piperno and Holst, 1998). Once domesticated, cassava cultivation started its expansion and finally being brought to the entire tropical America and the Caribbean region. Archaeological discoveries on remains on tools indicated that as early as 1800-900 B.C. cassava cultivation had been introduced to coastal Peru (Ugent et al., 1986). Cassava cultivation played an important role in the life and culture of pre-Columbian America. For example, its figure appeared on ceramic vessels of the Moche and Chimú cultures of Peru as shown in figure 1.5.



**Fig. 1.5 A vessel from Chimú culture (late stage of Moche culture, 1100-1470 AD) mimics the shape of a bundle of cassava roots.** Picture modified from (Ugent et al., 1986).

During the Age of Discovery cassava was introduced to other regions of the world by the explorers and navigators from the Europe (Cock, 1982). It is believed that cassava was brought to West Africa from Brazil by the Portuguese traders in the 16<sup>th</sup> century (Cock, 1982, Hillocks, 2002). Cassava was initially brought to the valley of Congo river (Aloys and Ming, 2006) and was not introduced to East Africa until 18<sup>th</sup> century when it entered via Madagascar and Zanzibar from overseas (Cock, 1982). Soon cassava started spreading inland through trading routes, but it was not until the early 20<sup>th</sup> century that cassava cultivation became a major source of food (Hillocks, 2002). For example, cassava had been brought to the Niger Delta by the Portuguese, but it did not replace yam as the Nigerien staple food until the local community was disrupted by the Spanish Influenza and the British expeditions in the early 20<sup>th</sup> century, when the rural communities were forced to seek for a more robust crop (Ohadike, 2009). In Asia, it was probably introduced to the Philippines and Goa by the Spanish and the Portuguese during the 17<sup>th</sup> or 18<sup>th</sup> century (Cock, 1982, Onwueme, 2002). In the late 19<sup>th</sup> century, cassava was cultivated widely in many parts of Asia. Nowadays, cassava cultivation is still mainly in the tropical



regions mentioned above, indicating their optimal conditions for cassava growth. Today the tropical regions between 30°N and 30°S of Africa, South America and Asia produce more than 70% of the total cassava yield (Jansson et al., 2009).

### 1.1.3 Importance of cassava in human life

The most utilised part of cassava is the storage root, which can be a staple food, an animal fodder, or an industrial raw material due to its high starch content (El-Sharkawy, 2006). Nowadays, cassava is the sixth most important crop world-wide, and its starchy roots and derived products are consumed as a major source of calories by 500 million to 700 million people (Olsen and Schaal, 2001, Wang et al., 2014). At the beginning of the new millennium, the global yield of cassava root reached 168 million tonnes per year (Jekayinfa and Olajide, 2007), whose order of magnitude is comparable to that of the more popular tuberous crop *potato* which is approximately 300 million tonnes per year (FAO, 2008). In 2013, Africa shared 57% of the annual yield, ranking the primary continent for cassava production. Africa was followed by Asia (31.9% global production) and America (11% global production), and Oceania contributed 0.1% of world's total cassava (FAOSTAT, 2013).

Cassava maintains a reliable yield on marginal and barren soil (Mkumbira et al., 2003), which makes it an ideal and popular choice for the subsistence of small-holder peasants in less developed regions. Methods of cassava processing and consumption vary depending on regions and cultivars. If the cultivar is 'sweet' with low cyanogen levels, the root can be eaten raw after peeling. However, studies (Wilson and Dufour, 2002) pointed out that in Africa and South America, cases are that cassava cultivars with higher cyanogenic compound levels ('bitter' cassava) are preferred due to its higher yield, which might be because the accumulated cyanide helped defend against pests.

Cassava has become an important part of the diet and culture in many cassava-growing regions. In Africa, due to a preference to 'bitter' cassava, processed cassava products are common (Mkumbira et al., 2003). Cassava roots are often boiled and/or fermented to remove the cyanogen, and processed to staple foods such as *gari* (cassava flour) or *fufu* (cassava dough) (Aloys and Ming, 2006, Jekayinfa and Olajide, 2007). In South America, a cassava flour (*farinha*)



similar to gari is prepared (Lancaster et al., 1982). Cassava can also be ground and roasted to make bread, such as the indigenous American flatbread *casabe* (Dufour, 2006). Cassava beer is also prepared (Barre, 1938). In tropical Asia, various cassava-based foods are also made, such as the fermented and fried cassava (named *tapai*) from Indonesia (Fathoni, personal communication).

Demands for a cleaner fuel from renewable resources is pushing cassava into the biofuel industry (Ziska et al., 2009). Biofuel is prepared from an abundant and low cost non-fossil biomass (Adelekan, 2010) and can partially replace fossil fuel. Currently the most widely used biofuel is ethanol from corn starch, but cassava starch as a source of bioethanol is drawing world's attention due to its cheaper price and greater yield (Zhang et al., 2003a). In the United States, corn ethanol production consumes more than 20% of its total corn yield, raising the price of maize for food purposes (Ziska et al., 2009). Cassava has ideal characters to replace maize in bioethanol production. Cassava plantations can be cultivated and harvested year round without breaks due to seasons, and the harvested cassava can be stored as dried chips for a continuous supply (Nguyen et al., 2007). Cassava provides a decent yield with limited water and fertiliser, which make it a crop of low input and cost. Finally, cassava root contains a high amount of starch (Jansson et al., 2009), which is favoured for ethanol fermentation.

#### **1.1.4 Why has cassava not been used wider?**

##### **1.1.4.1 Cyanogenic compounds**

Beside its favourable characters, cassava also suffers from several deficiencies that limited its utility. Cassava storage roots accumulate high levels of cyanogenic compounds, including linamarin (about 95%), the glycoside form of acetone cyanohydrin, and a small amount of lotaustralin (Santana et al., 2002, Siritunga and Sayre, 2003). Linamarin has the highest concentration in cassava leaves and root cortex, but relatively lower in root parenchyma (the 'flesh') (Cardoso et al., 2005). In the 'bitter' cultivars of cassava, it is the cyanogen that gives cassava the bitter taste. Although cultivars that contain less than 100 ppm of linamarin do exist (Bayoumi et al., 2010, Santana et al., 2002), in Africa and South America some cassava varieties have more than 500-1000 ppm of linamarin (Cardoso et al., 2005). Since

the maximal residual cyanogen recommended by the WHO is only 10 ppm, these cultivars are obviously harmful without appropriate processing.

The release of cyanide (HCN) is regulated by the enzyme linamarase. Linamarin is synthesised in leaves (through N-hydroxylation of valine and isoleucine) and distributed to the roots and stored in the vacuoles of cassava cells, but linamarase is located in the cell membrane and the inner cell wall (Mkpong et al., 1990, Sayre et al., 2011). Linamarase is more concentrated in cassava peels but it is also found in the parenchyma (Cooke et al., 1978). Linamarin and linamarase meet and react when cells are mechanically damaged, releasing acetone cyanohydrin. Acetone cyanohydrin decomposes to release HCN either by hydroxynitrile lyase or spontaneously when the pH is greater than 5 (Sayre et al., 2011). The decomposition of cyanogens discourages herbivores, but also provides a method to remove the toxin (Dufour, 2006). By mashing and fermenting, most of the cyanogens is converted to gaseous HCN and thus removed. Transgenic lines with linamarin synthesis blocked have been generated, but they do not grow well in the field due to impaired nitrogen transport ability via linamarin (Sayre et al., 2011).

Acute or chronic exposure to cassava cyanogen may cause diseases such as *konzo* (irreversible spastic paraparesis) and tropical ataxic neuropathy (TAN) (Cardoso et al., 2005, Siritunga and Sayre, 2003). Usually the cyanogenic compounds of cassava can be removed by soaking, boiling, or fermenting, but during a drought or food shortage, these methods may fail due to an overwhelmingly high cyanogen concentration or inadequate processing. The low protein content of cassava worsens the situation as sulphur-containing amino acids are involved in cyanide detoxification (Cliff et al., 2011, Ernesto et al., 2002).

#### **1.1.4.2 Low nutritional value of cassava**

Regardless of its high and reliable yield, the nutritional value of cassava root is relatively poor, due to cassava's low protein, vitamin, and mineral content (Nassar and Ortiz, 2010). Depending on the cultivar, protein varies from less than 1% to 5% of the dry weight of cassava roots, but mostly below 3% (El-Sharkawy, 2006, Stupak et al., 2006). To make it worse, some essential amino acids, such as lysine, leucine, and especially sulphur-containing amino acids, are particularly scarce in cassava. Compared to cereals, in which up to 14% of the dry weight is protein,

cassava is one of the staple foods with the lowest protein-energy ratio (Abhary et al., 2011). As a result, if a diet is mainly based on cassava, at least 1kg of dried cassava products must be consumed every day to achieve the British government's daily recommended protein intake (approximately 50g) (ONS, 2014).

Supplementing the diet with foods high in protein may improve the nutritional value of cassava, but they are likely to be unaffordable to many cassava growers.

Cassava leaves can be a feasible food supplement to cassava roots. When cassava is at the age of 11-12 months, 20 to 30% of the dry weight of the leaves is crude protein, with a well-balanced ratio of essential amino acids except for sulphur-containing amino acids (Eggum, 1970). Cassava leaves also require prolonged boiling to remove the cyanogenic compounds, which causes the loss of amino acids (Stupak et al., 2006).

#### **1.1.4.3 Diseases**

Cassava is susceptible to many diseases caused by microorganisms, among which cassava mosaic disease (CMD) and cassava brown streak disease (CBSD) are the most important.

Cassava mosaic disease (CMD) is the greatest epidemic attacking cassava in Africa (Sayre et al., 2011). CMD is caused by one or more of the three geminiviruses (African cassava mosaic virus, East African cassava mosaic virus, and Indian cassava mosaic virus) which are transferred by whiteflies (*Bemisia tabaci*) and infected stem cuttings (Zhang et al., 2005). The infected plants show characteristic leaf shape alteration and yellow-green mosaics on leaves, and suffer from slowed growth and impaired root formation (Bull et al., 2011). It is estimated that CMD causes up to 40% of the loss of cassava yield in Africa. During the 1930s to 1970s breeding work was done on producing CMD-resistant cassava lines through crossing with *Manihot glaziovii* Muell.-Arg. (Legg and Fauquet, 2004). Several resistant cultivars were bred and widely cultivated in Africa, but nowadays more advanced CMD-resistant cultivars are still required. Attempts to improve cassava's resistance to CMD by expressing antisense RNA in cassava (Bull et al., 2011, Sayre et al., 2011, Zhang et al., 2005), and RNA interference targeting CMV promoters or important genes has shown promising in improving the resistance and

recovery against CMD; however, a CMD-resistant cultivar of cassava has yet to be created through biotechnology (Vanderschuren et al., 2007b, Zhang et al., 2005).

Cassava brown streak disease (CBSD) is caused by cassava brown streak virus (CBSV, *Ipomovirus* sp.) which produces a brown corky necrosis in cassava storage roots (Alicai et al., 2007, Bull et al., 2011). CBSD is wide spread in East and the coastal Africa, and is probably transmitted by whiteflies (*Bemisia spp.*) or aphids (Hillocks and Jennings, 2003). The discoloration can be seen beneath the bark of young stems, as well as in the roots. The lesion may be constrained in tolerant cultivars but in more susceptible cultivars it may spread to the entire root. The necrosis often leads to a soft rot due to infection of secondary microorganisms. The disease can cause up to 70% of root weight loss depending on the sensitivity of the cultivar (Maruthi et al., 2005). Although CBSD may be controlled by field sanitation and using disease-free stem cuttings, CBSD-resistant cultivars of cassava are still yet to be created (Hillocks and Jennings, 2003).

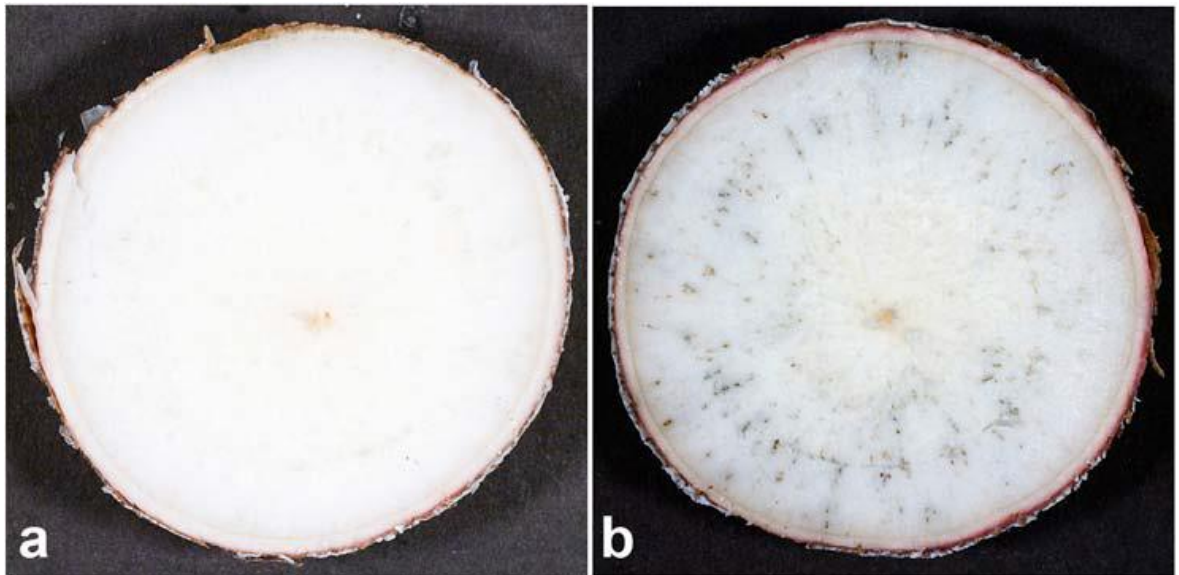
## **1.2 POST-HARVEST PHYSIOLOGICAL DETERIORATION (PPD)**

### **1.2.1 Post-harvest physiological deterioration and its importance**

Despite the deficiencies above, post-harvest physiological deterioration (PPD), the active physiological responses triggered by harvesting, poses the greatest hurdle in the utilisation of cassava.

PPD is triggered by the inevitable wounding that occurs during harvesting and handling, and is characterised by a rapid discolouration of cassava roots that occurs in 24-48 hours after harvest. Figure 1.6 shows cross-sections of the fresh and the deteriorated cassava storage roots. The discolouration appears first in the vascular system around the wounds (Reilly et al., 2007) from where it spreads to the rest of the root. In a few days after harvest, a blue-black colour will accumulate in the vascular bundles in the parenchyma (called *vascular streaking*) (Reilly et al., 2003). The deteriorated cassava roots have unfavourable appearance and taste, which makes them unacceptable in the market. In South America, sample cassava roots in markets are often cut open to show their freshness (Beeching, personal communication). As a result, once cassava roots have been harvested, there is only approximately three days for them to be transported and sold. The PPD

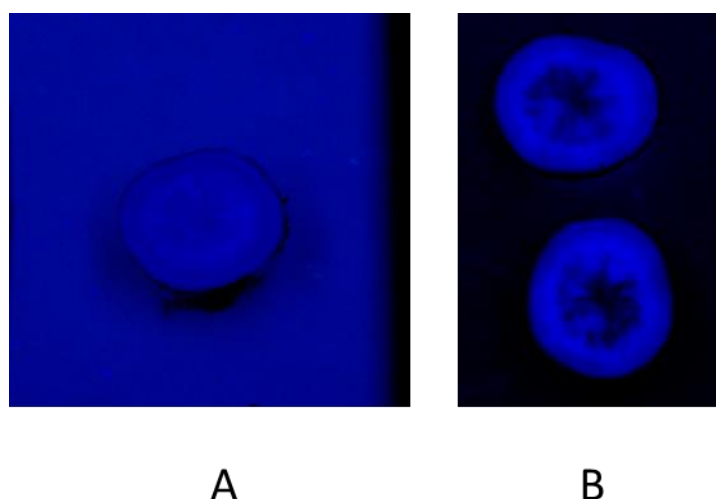
phenomenon makes cassava root a highly perishable commodity that cannot be delivered and marketed easily. It is estimated that in some of the major cassava producing countries, 10% to 20% of the cassava roots are lost due to the PPD (Ravi et al., 1996). Since many cassava producers are small-holder farmers in poor conditions, PPD makes it challenging for them to sell their products in the cash economy.



**Fig. 1.6 Cross-sections of fresh (a) and 48 hours post-harvest (b) cassava roots display the blue-black discoloration developed in vessels during the PPD response.** Photographs are of an unknown cultivar from a supermarket in UK. Photographs courtesy of John Beeching (University of Bath).

Many physiological responses are involved in the onset of PPD. The discolouration of vascular system involves blockage by tylose inclusions, which may possibly play a role in restricting potential pathogenic invasion (Rickard and Gahan, 1983). Respiration and consumption of oxygen increase, as well as the expression of peroxidation-related enzymes such as peroxidase or catalase. Bursts of reactive oxygen species (ROS) occur early after harvesting, which probably have multiple functions ranging from signals, defence and programmed cell death (Reilly et al., 2003). Synthesis of enzymes involved in the phenylpropanoid metabolism, such as phenylalanine ammonia-lyase (PAL), are up-regulated, accelerating the accumulation of phenolic secondary metabolites, which act as either biochemical

signals or defensive molecules (Wheatley and Schwabe, 1985). Among the phenolic metabolites, fluorescent hydroxycoumarins are in especial abundance, as the fresh cassava roots do not show significant fluorescence (figure 1.7 A) but the deteriorating roots develop a strong blue fluorescence under 366 nm ultra-violet light (figure 1.7 B). Increase of total cyanide, as well as the decomposition of starch to sugar, is also observed during the development of PPD (Uarota et al., 2014, van Oirschot et al., 2000). Although initially PPD is spontaneous without pathogenic activities, it is often followed by microorganism infection due to the disordered physiology, a process called secondary or microbial deterioration (Buschmann et al., 2000a).



**Fig. 1.7 Fresh (A) and 1 DAH (B) cassava samples emit fluorescence under 366nm UV light.** (A) Fresh cassava root does not show significantly visible fluorescence compared to filter paper beneath; (B) On 1 DAH cassava roots have developed significant fluorescence compared to the filter paper beneath.

## **1.2.2 Attempts to alleviate PPD**

### **1.2.2.1 Delaying the harvest**

Some methods have been developed to prolong cassava storage roots shelf-life by means of harvesting, handling, and storage, but most of them are not economical in terms of labour or monetary cost (van Oirschot et al., 2000). In many rural regions of Africa, mature cassava plants are allowed to stay ('live') in the field until needed

for local consumption (Reilly et al., 2007). Depending on the cultivar, cassava can be kept in field from months to a year after its normal harvesting time without significant quality loss. But this practice does not help with the marketing and transportation of cassava, and it ties up arable land that could be used to grow more crops.

#### **1.2.2.2 Pruning**

Pruning is an established way to delay PPD. The stems and leaves of a cassava plant are removed from 30 cm above the ground 10-30 days before harvest, which can delay PPD for more than five days (Tanaka et al., 1984). In the roots from pruned cassava plants expression of phenylpropanoid metabolism-related PAL (phenylalanine ammonia-lyase) and reactive oxygen species (ROS)-related peroxidase were reduced (they are normally up-regulated during PPD), indicating the roles for these enzymes in PPD (Tanaka et al., 1984). However, if a plant is pruned, its roots rapidly lose up to a quarter of their starch content (presumably due to the mobilisation of starch into sugar to enable re-growth), and the pruned roots are hardened as they become 'glassy', leading to unfavourable characters in cooking (Tanaka et al., 1984, van Oirschot et al., 2000). Due to these reasons, pruning has not become a successful commercial anti-PPD method.

#### **1.2.2.3 Exclusion of oxygen**

Another practiced method is to isolate the roots from the air (preventing oxygen access), that presumably prevents PPD through inhibiting the accumulation of reactive oxygen species (ROS). The most practical way is to coat individual roots through dipping them in paraffin wax once harvested (Reilly et al., 2003), which can seal the roots from the air and thus delay PPD long enough for them to be shipped trans-nationally. However, even though paraffin wax coating is a practical solution as long as the wax coating is not broken through further damage to the root, it is not economically viable for such a low value commodity on a local or national scale. It is only viable to national urban middle-class consumers or ones in developed countries who are prepared to pay over the odds for a traditional food.

#### **1.2.2.4 Curing**

Like other root crops, cassava roots may also be 'cured', which under optimal conditions leads to wound sealing and thus delayed PPD. A meristem and wound cork can form slowly beneath the wounds under a relative humidity of 80-85%, and this process is accelerated at a higher temperature (up to 40 °C) but with an increased risk of microbial infection (Ravi et al., 1996). Cured cassava roots develop PPD slower, with less discolouration and weight loss than normally handled roots; however, the tricky demands and limited benefits make the curing method uneconomical. In addition, once the cured roots are wounded again, a rapid PPD develops despite the previous treatment.

#### **1.2.3 The mechanism of PPD**

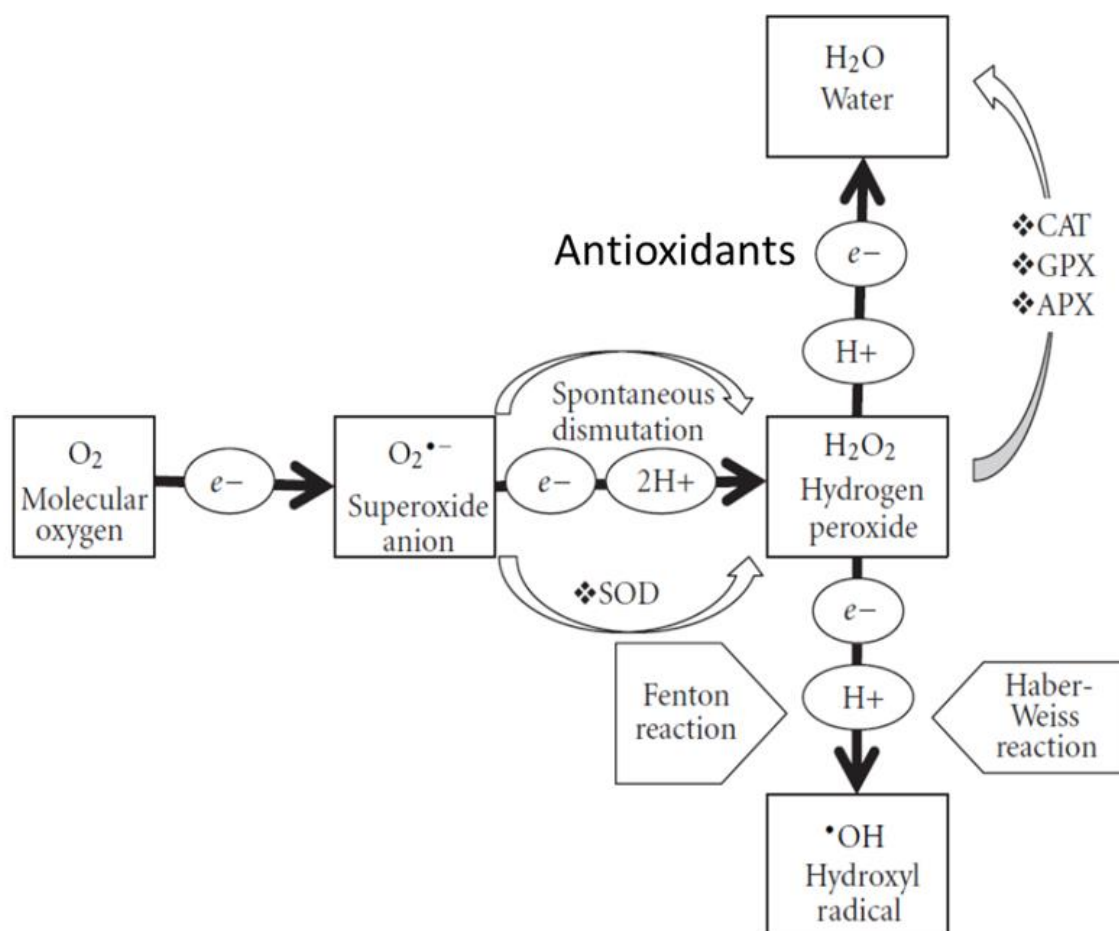
Due to its importance, much research has been done on the mechanism of PPD (Blagbrough et al., 2010, Buschmann et al., 2000b, Reilly et al., 2003, Sayre et al., 2011, Wheatley and Schwabe, 1985). The initiation of PPD is active, as protein translational inhibition by cycloheximide delays its development (Beeching et al., 1998). The early stages of PPD involve the release of ethylene, the alteration of the composition of phospholipids, increased respiration rate, accumulation of reactive oxygen species (ROS, e.g. H<sub>2</sub>O<sub>2</sub>), and the increased synthesis of the enzymes regulating the events above, such as catalase, peroxidase, and superoxide dismutase (Lalaguna and Agudo, 1989, Sayre et al., 2011, Tanaka et al., 1984, Zidenga et al., 2012). In addition, phenolic metabolites also accumulate in the cassava roots, being accompanied by the increased activities of the related enzymes, such as PAL and polyphenol oxidase (Buschmann et al., 2000b, Reilly et al., 2003).

Unwounded cassava roots generate no detectable ethylene emission and ethylene production starts a few hours after wounding of the roots (Hirose, 1986, Plumbley et al., 1981). In some tuberous crops such as sweet potato, the wound-induced ethylene enhances the activity of peroxidase, indicating a role in the wound response (Imaseki, 1970, Imaseki et al., 1968). These observations led to the suggestion that ethylene triggered a cyanide-insensitive respiration pathway which led to the activation of peroxidase (Solomos and Laties, 1976). However, the



application of exogenous ethylene to fresh cassava root samples did not lead to a significant increase in discoloration (Hirose, 1986). But application of exogenous ethylene did increase the expression of peroxidase genes but that of other ROS-related genes was not significantly altered (Reilly et al., 2003). Therefore, while ethylene does play a role in coordinating more general wound responses, its specific role in PPD remains unclear.

Oxidisation-related processes appear to be critical to the development of PPD since the most successful commercial way of delaying PPD is to prevent oxygen access to the roots by coating them with wax (Buschmann et al., 2000b, Reilly et al., 2003). On the onset of PPD, an oxidative burst is triggered within 15 minutes of harvesting, producing and accumulating large amount of ROS, such as superoxide and  $\text{H}_2\text{O}_2$ , as a part of the stress response not only for pathogen defence but also signalling in further responses (Buschmann et al., 2000a, Wheatley and Schwabe, 1985, Wojtaszek, 1997). The generation and scavenging of ROS are shown in figure 1.8. The major reactive oxygen species are superoxide ion ( $\text{O}_2^{\cdot-}$ ), hydroxyl radical and hydrogen peroxide that are generated from the electrons leaking out from the electron transport chains (Xu et al., 2013). Generally, the oxygen molecule is converted to superoxide radicals, which are processed to less harmful  $\text{H}_2\text{O}_2$  catalysed by SOD. The  $\text{H}_2\text{O}_2$  is decomposed to  $\text{H}_2\text{O}$  and  $\text{O}_2$  molecule by catalase (CAT) or ascorbate peroxide (APX) with the oxidation of relevant anti-oxidants (Owiti et al., 2011). While oxidative bursts are associated with more general wound responses in plants, it is possible that in cassava PPD there is a relationship with the cyanogenic potential of the roots (Zidenga et al., 2012). In a low-cyanogen line of cassava, the  $\text{H}_2\text{O}_2$  generation is significantly lower than the high-cyanogen lines, and externally applied cyanide promotes the production of ROS. These findings can be explained as follows: wounding disrupts cassava cells, releasing cyanide from the cyanogenic compounds (e.g. linamarin); cyanide inhibits the cytochrome c oxidase in the electron transfer chain in mitochondria (Cooper and Brown, 2008), forcing excessive reducing power to ROS generation. Supporting evidence comes from expressing cyanide-insensitive alternative oxidase (AOX) in cassava, which led to delayed ROS accumulation and PPD discoloration (Zidenga et al., 2012).



**Fig. 1.8 The metabolism of the generation and removal of plant reactive oxygen species (ROS).** Picture modified from (Sharma et al., 2012).

While ROS is harmful to plant cells, this is dependent on cellular levels and the species in question; for example, hydrogen peroxide is involved in hydroxyproline-rich glycoprotein insolubilisation and lignin formation, and at higher levels they play roles in defence. However, plants possess regulatory mechanisms to control their activity so as to maintain homeostasis. In cassava landraces with high content of carotene, which is an anti-oxidant, PPD develops slower, supporting the argument that a reducing environment alleviates the damage by oxidation (Chavez et al., 2005). Gene expression analysis using cDNA microarray showed that 18% of the genes activated during PPD were related to the turnover and detoxification of ROS, and these included genes for catalases (CAT, EC 1.11.1.6), ascorbate peroxidases (APX, EC 1.11.1.11) and peroxidases (POX, EC 1.11.1.7) (Reilly et al., 2007). A study of changing protein accumulation during PPD

using isobaric tags for relative and absolute quantification (iTRAQ) found that the accumulation of superoxide dismutase (SOD), catalase (CAT), and the regulatory protein nucleoside diphosphate kinase (NDK) of CAT, was increased during PPD development (Owiti et al., 2011). In cassava roots, the ROS scavenging mechanism may be inadequate to stop ROS from causing damage, due to the accumulation of the detoxifying proteins after 24 hours by which time the response has initiated. Xu et al. succeeded in delaying PPD development in cassava roots for up to 10 days after harvest through the co-overexpression of CAT and SOD in cassava, which suggests that an enhanced ROS removal can help keep the PPD in control (Xu et al., 2013). Reducing storage proteins are normally accumulated in many tuberous crops during development but not in cassava roots, which could explain why they are particularly sensitive to the oxidative damage immediately on harvest (Stupak et al., 2006).

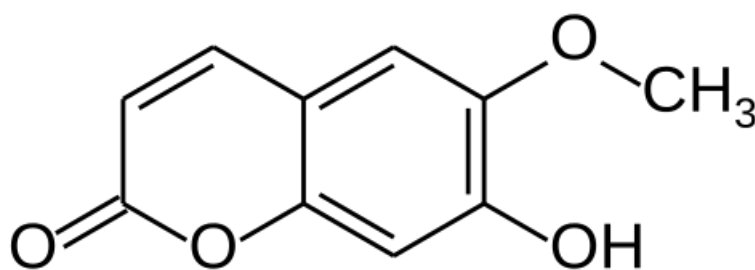
Another important component of the abiotic PPD responses is the accumulation of phenolic compounds. During the development of PPD, the rapid discoloration is characterised by increased levels of phenolic compounds in the root (Rickard, 1985). The most significantly accumulated phenolic compounds were hydroxycoumarins, especially scopoletin and scopolin (Wheatley and Schwabe, 1985), but many other phenolic compounds, including esculin, proanthocyanidins, catechin, and gallic acid, have also been identified to be related to the PPD development (Sayre et al., 2011). In general, the scopoletin level in cassava root increases up to two hundred fold from a relatively low level (less than 1 ng/mg) in 1-2 days after harvest. In PPD development, these phenolic compounds may act as anti-microbial and anti-oxidant reagents to keep the over-active oxidative responses and potential pathogenic invasion in control (Buschmann et al., 2000b). Some phenolic compounds, especially flavonoids, are able to inhibit the growth of microorganisms, probably by interfering with the activities of bacterial enzymes and proteins (Cowan, 1999, Rauha et al., 2000). Phenolic compounds are able to donate electrons and thus scavenge the ROS with their hydroxyl groups (Balasundram et al., 2006), and in many plants the anti-oxidant capacity is proportional to the total phenolic content (Cai et al., 2004). To support this, certain kinds of phenolic compounds, especially scopoletin, form a blue-black precipitation similar to the discoloration in cassava roots if mixed with H<sub>2</sub>O<sub>2</sub> and a crude extract of cassava roots that are undergoing the PPD process (Beeching et al., 2002). A similar result is achieved if horse-radish peroxidase is used instead of cassava

extract, indicating that scopoletin and other phenolic compounds can serve as anti-oxidants to scavenge the endogenous ROS. As a summary, the oxidization of phenolic compounds, including scopoletin, scopolin, esculetin and proanthocyanidins, alleviates oxidative stress while contributes to the discoloration (Xu et al., 2013).

#### **1.2.4 Scopoletin and phenolic secondary metabolites in PPD**

As discussed above, the accumulation of stress response metabolites is an important link in the process of PPD development. A wide range of compounds are involved in PPD development, including not only coumaric and phenolic compounds but also other classes. For example, phytosterols and fatty acids (e.g. palmitic acid, linoleic acid, and oleic acid), as well as their derivatives, accumulate during the PPD response (Sakai et al., 1986). The accumulation of a large family of 22 different diterpenic compounds was also identified in wounded cassava roots (Sakai and Nakagawa, 1988), which is unusual in plant stress responses. However, it is hydroxycoumarins that accumulate most 'dramatically' after harvest during PPD (Bayoumi et al., 2010). The most significant one of these hydroxycoumarins are scopoletin and its glucoside scopolin, whilst esculetin and its glucoside esculin also accumulate, but to less significant quantities. Other PPD related metabolites include leucoanthocyanins, delphinidin, flavan-3-ols, catechin, and gallocatechin (Buschmann et al., 2000b, Sakai et al., 1994, Sakai et al., 1986). These secondary metabolites may act as anti-oxidants or antimicrobial agents (Buschmann et al., 2000b, Sakai and Nakagawa, 1988).

Of these metabolites, scopoletin is of particular interest as it shows the most dramatic accumulation during the PPD. Scopoletin (or 7-hydroxy-6-methoxychromen-2-one) is a hydroxycoumarin compound synthesised as part of phenylpropanoid metabolism, and its structure is shown in figure 1.9. There is evidence showing that scopoletin is involved in plant defensive responses in cassava, tobacco, tomato, and other plants (Sudha and Ravishankar, 2002, Sun et al., 2014).



**Fig. 1.9 The chemical structure of a scopoletin molecule.**

At the onset of PPD, the level of scopoletin in cassava roots rises rapidly, increasing from less than 1 ng/mg to 60-80 ng/mg in fresh roots in two days, and remains at a high concentration during the further development of PPD (Buschmann et al., 2000b). The dramatic accumulation of scopoletin in PPD indicates that it plays an important role in the deterioration process. Supporting this, a pruning treatment that delays PPD also lowers scopoletin synthesis in cassava roots (van Oirschot et al., 2000). In tobacco, scopoletin acts as one of the major  $\text{H}_2\text{O}_2$  scavengers after the wound-triggered oxidative burst, and contributes in the formation of a lignin-like substance to seal the wounds (Dorey, 1999). Scopoletin may scavenge ROS with its hydroxyl group; to support this, scopolin, the glycoside of scopoletin whose hydroxyl group is protected by a glucose residual, does not show ROS scavenging activity (Reilly et al., 2003). Wheatley et al. applied a range of phenolic compounds to fresh cassava root samples, and found that scopoletin was the only one that induced a rapid PPD discoloration, implying that scopoletin might act as a signalling molecule to trigger PPD (Wheatley and Schwabe, 1985). However, there is no other evidence in the literature suggesting that scopoletin might have a signalling function, while there is considerable evidence for its anti-oxidant and anti-microbial action (Ojala et al., 2000, Shaw et al., 2003, Shukla et al., 1999). In addition, peroxidase has a high affinity to scopoletin, oxidising it to a coloured insoluble metabolite (Edwards et al., 1997, Wheatley and Schwabe, 1985), indicating scopoletin as a source of the discolouration observed during PPD. Further evidence supporting scopoletin's role in the discoloration is that at the earliest stages of PPD, fluorescent material accumulated firstly around the xylem vessels (Buschmann et al., 2000b), since the blue-black discoloration and tylose formation start from the vascular system, it is likely that scopoletin is involved in the formation of discoloration and vascular blockade. In sunflower (*Helianthus annuus*)

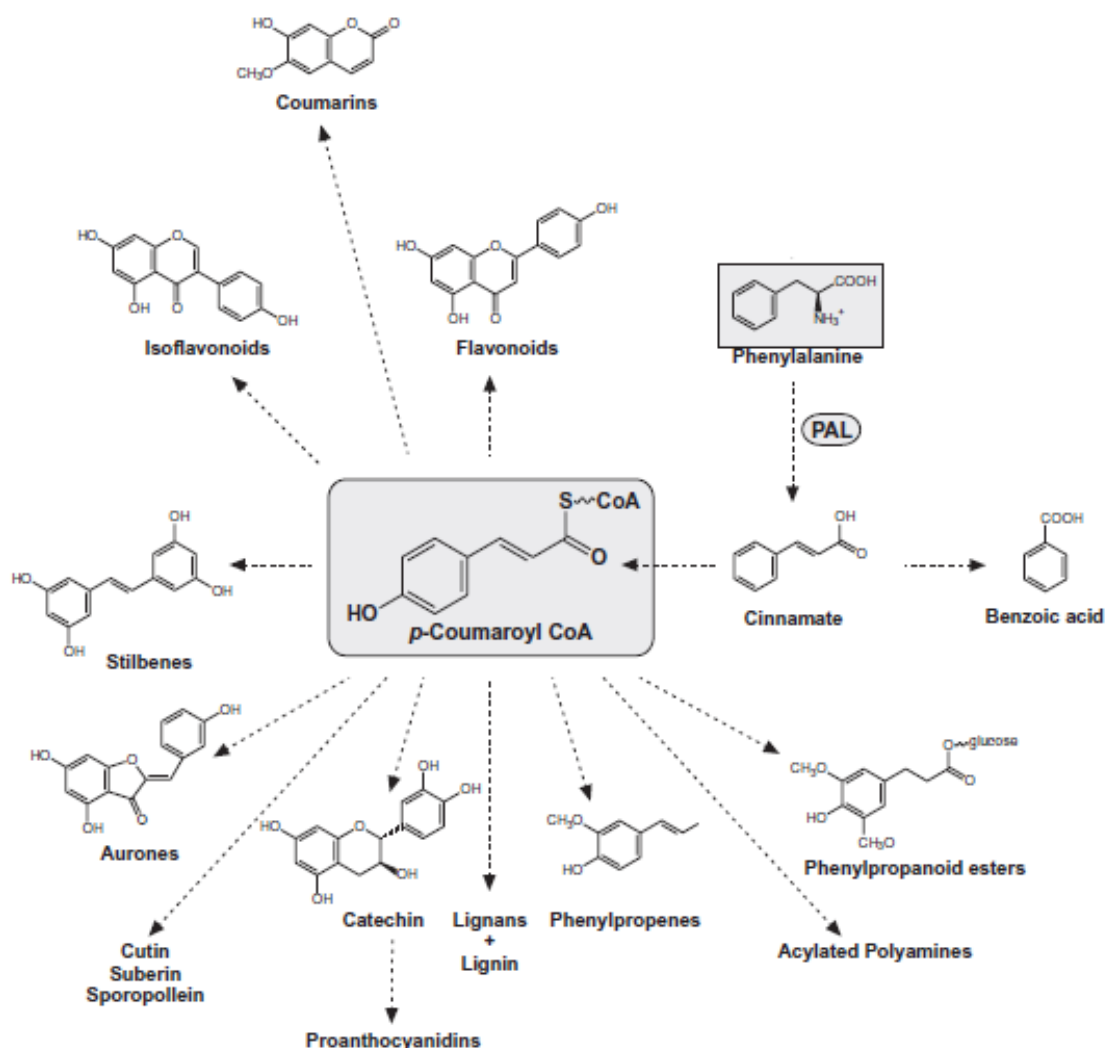
a detached leaf or an extract of sunflower peroxidase incubated with scopoletin and  $\text{H}_2\text{O}_2$  resulted in an insoluble blue product; however, if  $\text{H}_2\text{O}_2$  was absent, or catalase is provided, there was no blue product formation (Edwards et al., 1997, Gutierrez, 1995). The conclusion from these data is that scopoletin is an important factor in the discoloration of cassava that occurs during the PPD process.

### **1.3 SCOPOLETIN SYNTHESIS AND PHENYLPROPANOID METABOLISM**

#### **1.3.1 Phenylpropanoid metabolism in higher plants**

Phenylpropanoid metabolism is conserved in higher plants, and its products play important roles in structural support (lignin), water transport, UV protection (flavonoid pigments), reproduction, and tolerance to biotic and abiotic stresses (Vogt, 2010). The key classes of products of phenylpropanoid metabolism are shown in figure 1.10. There is evidence that at the time when plants colonised land, the ancestor of terrestrial plants acquired PAL, the key entry enzyme to phenylpropanoid metabolism, from a bacterial genome through horizontal gene transfer (Emiliani et al., 2009). Subsequently, other enzymes necessary for the diverse aspects of phenylpropanoid metabolism evolved.

Products of phenylpropanoid metabolism are involved in many responses to biotic or abiotic stresses, as well as in many other important processes, such as cell-wall formation and biomass accumulation (Cass et al., 2015, Vogt, 2010). The mechanism and genes involved had been studied extensively in model plants such as *Arabidopsis thaliana* via T-DNA gene knock-out mutants (Kai et al., 2006, Ossowski et al., 2008), enabling us to understand the functioning of this complex network of pathways.



**Fig. 1.10 Phenylpropanoid metabolisms produce many important secondary metabolites, including scopoletin and other coumarins.** Picture modified from (Vogt, 2010).

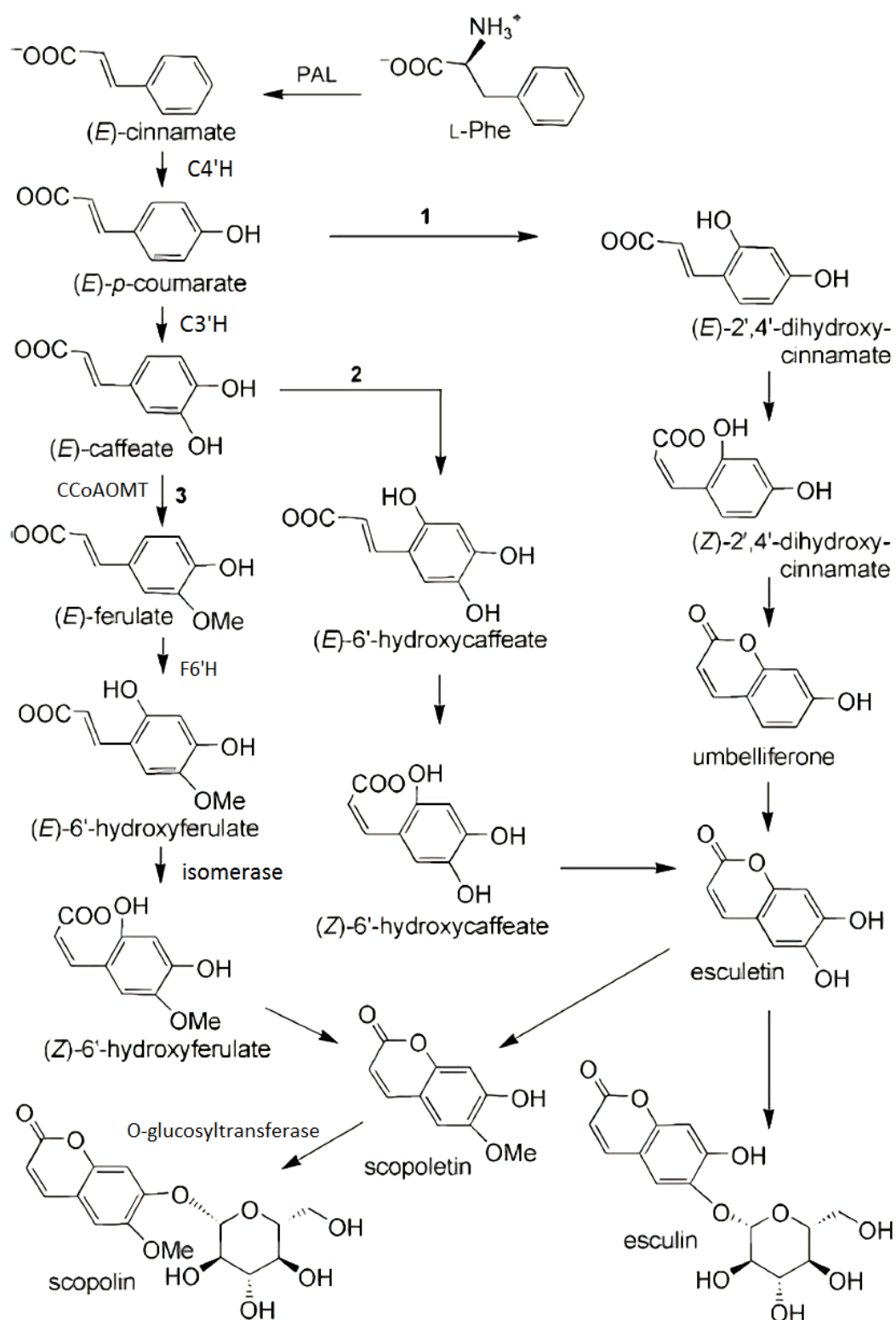
The entry point of phenylpropanoid metabolism is the amino acid L-phenylalanine (Phe). Phe is converted to trans-cinnamate by a tetrameric enzyme phenylalanine ammonia-lyase (PAL), thereby diverting the amino acid from primary to secondary metabolism (Cass et al., 2015). Cinnamate can be converted to p-coumarate via 4'-hydroxylation by cinnamate 4'-hydroxylase (C4'H) (Kai et al., 2006). Studies show PAL also has activity towards tyrosine, converting it to p-coumarate. Coumarate is linked to co-enzyme A (CoA) by 4'-coumarate co-enzyme A ligase (4'CL), becoming p-coumaroyl-CoA and thus activated, enters the diversity of phenylpropanoid metabolism, being further converted into many other phenolic metabolites through various pathways (Bayoumi et al., 2008a, Kai et al., 2006, Tian

et al., 2013). 3-Malonyl CoA and p-coumaroyl CoA can form tetrahydroxychalcone or trihydroxychalcone through the action of chalcone synthase (CHS) and chalcone reductase (CHR), and these intermediates are in turn processed to a wide series of flavonoid compounds via different pathways, including anthocyanins, which are the pigments giving flowers and fruits their colours (Emiliani et al., 2009, Winkel-Shirley, 2002a). Otherwise p-coumaroyl-CoA is converted to caffeoyl-CoA through 3'-hydroxylation, which is the starting point of the synthesis of lignin monomer and scopoletin (Kai et al., 2008). Lignin is an important product of the phenylpropanoid metabolism, being not only a crucial polymer in cell-wall formation, but also the destination of about 30% of the annually fixed carbon dioxide (Humphreys and Chapple, 2002).

### **1.3.2 Scopoletin biosynthesis involves three pathways in phenylpropanoid metabolism**

In cassava, scopoletin biosynthetic metabolism is part of the lignin biosynthesis. With the help of isotope labelling technology, Bayoumi et al. explored three possible metabolic pathways that led to scopoletin and its inert storage glycoside scopolin, as shown in figure 1.11. Among the three alternative pathways, the one converting phenylalanine to scopoletin via ferulate (pathway 3, figure 1.11) proved to be the major one producing 90% of the total scopoletin in cassava (Bayoumi et al., 2008a).





**Fig. 1.11 Three alternative pathways are found in the biosynthesis of scopoletin in cassava.** Number 1, 2, and 3 indicate the entry points of three possible pathways. The enzymes involved in pathway 3 are indicated around the arrowheads. CoA is omitted from the structure of the compounds. (Picture modified from Bayoumi et al. 2008b)

In the major pathway (pathway 3), phenylalanine (Phe) is first processed to cinnamate and then p-coumarate as per above, and then its carboxyl group is linked to a coenzyme A (CoA) through a thioester bond by 4'CL (Ferrer et al., 2008). The activated p-coumaroyl CoA is converted to caffeoyl CoA by adding a hydroxyl group to its 3'-carbon by C3'H. Then the 3'-hydroxyl group is methylated by caffeoyl CoA O-methyltransferase (CCoAOMT), becoming feruloyl CoA, the coenzyme A-linked form of ferulate. To provide the lactone structure of scopoletin, a 6'-hydroxyl group needs to be added to the benzene circle of feruloyl CoA, a step that requires a 2-oxoglutarate-dependent dioxygenase (2OGD) (Kai et al., 2008). This enzyme is feruloyl CoA 6'-hydroxylase 1, known as F6'H1 in *Arabidopsis thaliana*. Through the use of T-DNA insertion mutants, it has been shown that in *A. thaliana*, F6'H1 is the only gene regulating the formation of 6'-hydroxyferuloyl CoA, and if knocked out, downstream products such as scopoletin do not accumulate (Kai et al., 2008). For lactonisation, the 6'-hydroxyl and carboxyl groups need to be converted to the Z-isomer via E-to-Z isomerisation by an isomerase (Gnonlonfin et al., 2011). (Z)-6'-hydroxyferuloyl CoA can form the lactone structure spontaneously, generating a scopoletin molecule. Scopoletin can be further glycosylated to its inactive glycoside, scopolin, for storage in the vacuole by an O-glucosyltransferase. Scopolin can be deglycosylated to release scopoletin by the action of a  $\beta$ -glucosidase (Ahn et al., 2010).

Alternatively, scopoletin can also be synthesized through esculetin. P-coumaroyl CoA is 2'-hydroxylated to form 2', 4'-dihydroxycinnamate, then further processed to esculetin through a series of reactions, and finally methylated to form scopoletin (pathway 1). Another possible pathway branches from caffeoyl CoA, which is 6'-hydroxylated and converted to esculetin and then scopoletin as per above (pathway 2). However, these two alternative pathways between them supply only less than 10% of the total scopoletin synthesised in deteriorating cassava roots (Bayoumi et al., 2008a, Bayoumi et al., 2008b).

## **1.4 STRESS RESPONSES IN OTHER ROOT AND TUBEROUS CROPS**

### **1.4.1 Potato (*Solanum tuberosum*)**

Compared to cassava, many other root (or tuber) crops recover better from wounding and stresses, making their storage easier. Critical comparisons between the post-harvest responses of other root crops with that of cassava may provide useful insights into PPD.

Dehydration and infection are the major stresses caused by wounding. Potato tubers can be wounded during harvest, which triggers the accumulation of polyphenolic and polyaliphatic compounds in cell wall that suberize and seal the wounds, preventing desiccation and microorganism invasion. This is followed by the formation of a layer of wound periderm beneath the suberized cells (Lulai et al., 2008). Wounding triggers massive ethylene production in potato tubers, but inhibition of ethylene synthesis and signalling hardly alters the accumulation of polyphenolic compounds. Therefore, the role of ethylene in potato stress responses is still controversial (Lulai and Suttle, 2004, Reyes and Cisneros-Zevallos, 2003). However, synthesis of abscisic acid (ABA) is also activated by the wounding, and a down-regulated ABA production delays the accumulation of polyphenols and polyaliphates, which are important components of suberin formation (Franke and Schreiber, 2007). Regardless of the mechanisms behind these events, stimuli such as infection trigger the rapid increase of PAL gene expression, leading to increased synthesis of phenolic compounds produced by phenylpropanoid metabolism (Hahlbrock and Scheel, 1989).

### **1.4.2 Sweet potato (*Ipomoea batatas*)**

At ambient temperature, sweet potato can be stored from days to weeks depending on cultivars and storage conditions (Ray and Ravi, 2005). Nowadays sweet potatoes may be stored for up to 10 months in warehouses with controlled temperature in the US (Holmes and Stange, 2002, Ravi et al., 1996).

The skin of sweet potato is thin and vulnerable, which makes it likely to be damaged during harvest and handling. Wounding triggers the synthesis of ethylene in sweet potato roots, and ethylene will in turn activate metabolic responses to the

injuries, including the up-regulation of PAL and peroxidase activities (Imaseki et al., 1968). Phenolic compounds accumulate in the phellem, phloem and the tissue beneath the periderm immediately after wounding (Schadel and Walter, 1981). The up-regulation of peroxidase activity indicates ROS is involved in the wound response mechanisms, and the responses induced by ethylene are known to increase the resistance to fungal infection (Stahmann et al., 1966). Although PAL, the entry enzyme to the phenylpropanoid metabolism, is also involved, sweet potato does not develop discolouration during abiotic wounding-responses (Stahmann et al., 1966). After harvest, the roots can seal their wounds and harden the skin in a *curing* process, and the cured roots can be stored with reduced losses (Ray and Ravi, 2005). The curing process causes the exposed parenchyma to be suberized, forming a periderm beneath the wounds to protect the roots (Ray and Ravi, 2005).

#### **1.4.3 Yam (*Dioscorea* spp.)**

Yams are many species of tuber crops from the family Dioscoreaceae. The cortical periderm of yam is thin and vulnerable, which is often wounded during handling (Ejechi and Souzey, 1999). Yam tubers are propagating organs and have evolved to be able to cure their wounds and remain alive during dormancy (Ravi et al., 1996). A rapid response is activated after wounding, and a cork periderm will form in five days beneath the wounds (Ravi et al., 1996). Abscissic acid (ABA) is found to induce the expression of PAL and the accumulation of phenolic compounds (Cottle and Kolattukudy, 1982). Yam tubers start accumulating polyphenols when wounded, and at the fifth day after wounding, there will be 5-6 times the polyphenols present in fresh tubers (Ikediobi et al., 1989). Wounding also leads to the expression of oxidation-related enzymes (polyphenol oxidase, peroxidase, lipoxygenase, etc.), but polyphenol oxidase has the most significant up-regulation, indicating the role of polyphenols in anti-oxidation and stress resistance (Ikediobi et al., 1989). Therefore, the formation of polyphenol is related to the suberisation of wounds (Rhodes and Woollorton, 1978).

#### **1.4.4 Taro (*Colocasia esculenta*)**

Taro is an Araceae (aroid) crop planted for its starchy corms in Asia-Pacific and Africa. Taro corm, like many other tuberous crops, is soft and moist, rendering it vulnerable to mechanical and biotic stresses (Ravi et al., 1996). Taro corms are able to heal wounds caused during harvesting and handling, but high temperature and humidity are required. These corms maintain a high respiration rate under high temperature after harvesting, but a successful wound healing also requires a high storage temperature (Agbor-Egbe and Rickard, 1991, Ravi and Aked, 1996). Under external stresses, phenolic accumulation is triggered in taro as a part of the defensive response (Sahoo et al., 2009). However, during the storage of taro, phenolic level in the corms will gradually decrease (Medoua et al., 2007). During six months storage, the levels of phenol oxidase and peroxidase activities in taro were maintained steadily, indicating homeostatic and mild post-harvest responses (Ohazurike and Arinze, 1996).

#### **1.4.5 What can we learn?**

In general, plant wounding response involves a signalling network, in which jasmonic acid and ethylene play important roles (Reyes and Cisneros-Zevallos, 2003, Schilmiller and Howe, 2005). During the wounding response of storage roots and tubers, phenolic compounds accumulate as anti-oxidants and phytoalexins. However, the hundred fold accumulation of certain phenolic compounds (such as scopoletin) in cassava is unusual among tuberous crops (Reyes and Cisneros-Zevallos, 2003).

Unlike many other tuberous crops, the cassava storage root is not a propagule; rather it serves only as a storage organ. Since the storage root serves no biological function once detached from the plant, cassava does not 'need' the ability to heal wounds or restore homeostasis under such circumstances. As a result, detached cassava roots appear to be unable to regulate the spontaneous responses to wounding and stresses; thereby, failing to prevent the reactions from cascading out of control. In other tuberous crops, the wound response reactions are followed by wound sealing that leads to the inhibition of the signalling system, preventing unnecessary over-reactions (Buschmann et al., 2000b). However, in cassava the

wound healing is inadequate and the wound response signalling spreads throughout the root without effective down-regulation (Reilly et al., 2003). PPD is the result of this over-reaction. Cassava can also suffer from microbial infections post the initial PPD reaction, but the rapid onset of PPD makes it less affected by the biotic rotting (Mwenje et al., 1998). It is pertinent that while detached cassava roots fail to repair wounds, if they remain attached to the plant as a whole, their wounds can heal (Buschmann et al., 2000b, Mwenje et al., 1998). This indicates that although cassava storage roots do possess the necessary mechanisms for effective wound repair and return to homeostasis, these do not function properly when detached by harvesting. Although various methods have been developed to prevent or delay PPD, none of these is generally applicable for such a low value commodity. However, biotechnological approaches may provide further insights into PPD responses and may indicate potential solutions to the problem.

## **1.5 RESEARCH STRATEGY**

In view of the key role that the dramatic accumulation of scopoletin and its oxidation appear to play in the development of PPD in cassava, this research project aims to further dissect scopoletin's role in the response. To this end the overall goal of the project is to block the synthesis of scopoletin in cassava through the inhibition of a key step in that process; thereby, enabling the comparison of cassava plants with decreased scopoletin content to those with normal synthesis in terms of their secondary metabolite profile and PPD response. The proposed strategy to address this question is to: identify cassava candidate genes that are functionally similar to *Arabidopsis* feruloyl CoA 6'-hydroxylase 1 (F6'H1); confirm the functional identity of these candidate genes through their ability to complement *Arabidopsis* T-DNA insert mutants of F6'H1; produces transgenic cassava containing RNAi constructs with the potential to inhibit the identified cassava F6'H1 candidate genes; characterise single-insert lines for this construct in terms of their general phenotype, secondary metabolite profile and PPD response; and synthesise these results into a model of PPD responses in cassava.

## 2 INSTRUMENTS, MATERIALS, AND METHODS

### 2.1 INSTRUMENTS AND CHEMICALS

The chemicals and reagents used in this study were obtained from Sigma-Aldrich unless otherwise stated. The plastic consumables used were obtained from Fisher Scientific unless otherwise stated. PCR and other reactions set up in 0.2 mL tubes that require incubation were carried out in a Thermo Cycler PTC-200 (MJ Research). DNA and RNA samples electrophoresed on gels were visualized in a GDS 7500 UV transilluminator (UVP) with a Grab-IT 2.0 (Synoptics Ltd.) imaging system. Centrifugations were run in different centrifuges depending on the volume and speed requirement of the sample. A Universal 32R centrifuge (Hettich) was used for low-speed centrifugation (e.g. for cell precipitation); an Allegra 25R centrifuge (Beckman Coulter) was used for high speed centrifugation (e.g. for DNA precipitation); and a MiniSpin microfuge (Eppendorf) was used to centrifuge small amount (less than 2 mL) of samples. Bacterial cultures that required shaking incubation were carried out in an Orbital Incubator (Sanyo) and the Petri dishes were incubated in a Function Line microbiological incubator (Heraeus). *Arabidopsis thaliana* plants and *in vitro* cassava were grown in a Fitotron plant growth chamber (Weiss Gallenkamp, also referred to as a 'controlled environment suite'). The electro-transformation of *Agrobacterium* was carried out through electroporating with a MicroPulser electroporator (Bio-Rad). The absorbance of nucleic acid or cell culture samples was measured with a Helios spectrophotometer (Thermo Scientific, for large quantity of samples), or a NanoVue spectrophotometer (GE, for microliters of nucleic acid samples).

### 2.2 PLANT MATERIALS

Wild-type *Arabidopsis thaliana* (ecotype Columbia-0, referred to as *Arabidopsis* below) seeds were obtained from Dr Julia Tratt (University of Bath, Bath). The *Arabidopsis* F6'H1 T-DNA insert mutant lines (Salk\_129938, Salk\_005232 and Sail\_1252\_A10) were obtained from the Nottingham *Arabidopsis* Stock Centre (NASC). Wild-type cassavas (cultivar: tropical *Manihot* series (TMS) 60444) were obtained from the glasshouse of Department of Biology and Biochemistry, University of Bath.

## 2.3 BACTERIA STRAINS

10-beta Competent *E. coli* (High Efficiency) (New England Biolabs) were used in *Escherichia coli* (referred to as *E. coli* below) transformation. Homemade electro-competent *Agrobacterium tumefaciens* (referred to as *Agrobacterium* below) was prepared from the *Agrobacterium* line GV3101 (with disarmed pGV3101 Ti plasmid) strain (Koncz and Schell, 1986), which was kindly provided by Dr Baoxiu Qi (University of Bath, Bath). *Agrobacterium* line LBA4404 (with disarmed pAL4404 Ti plasmid) (Hoekema et al., 1983, Negrotto et al., 2000) was provided by ETH-Zurich. The details of *Agrobacterium* lines are shown in table 2.1.

Names of lines	Ti plasmids	Genomic antibiotics	Ti plasmid antibiotics
<b>GV3101</b>	pGV3101	Rifampicin	Gentamicin
<b>LBA4404</b>	pAL4404	Rifampicin	Tetracycline

**Table 2.1 Details of *Agrobacterium tumefaciens* lines used in the study.**

## 2.4 PLASMIDS

The expression vectors (for *Arabidopsis* F6'H1 insert mutants) and RNAi vectors (for wild-type cassava) were based on pCAMBIA 1305.1 vector from the Centre for the Application of Molecular Biology to International Agriculture (CAMBIA, <http://www.cambia.org>, GenBank accession: AF354045). The vector contains the binary borders for *Agrobacterium*-mediated transformation, a hygromycin resistant gene (hptII) to select transformed plants, and a kanamycin resistant gene to select bacteria that carry it. A constitutive cauliflower mosaic 35S promoter (CaMV 35S) regulates the expression of the T-DNA region (a GusPlus reporter gene). The pCAMBIA 1305.1 plasmid was modified to generate two Gateway<sup>®</sup>-competent vectors with different promoters to be used in Gateway<sup>®</sup> cloning. The CaMV 35S-GusPlus structure was replaced with either a Gateway<sup>®</sup> cassette (Invitrogen) under the control of a constitutive CaMV 35S promoter (done by Dr Andrew Preston and Dr Jerry King, University of Bath), or a Gateway<sup>®</sup> cassette (Invitrogen) under the control of a root-specific StPAT promoter from the potato storage protein PATATIN (done by Dr Simon Bull, University of Bath).



pDONR™ /Zeo (Invitrogen) vector was used to shuttle a cloned DNA fragments, which were flanked by attB sites, to the final destination, i.e. the expression vectors and the RNAi vectors. In the pDONR™/Zeo vector, there are a Zeocin resistant gene, and a lethal gene *ccdB*, which stops the *E. coli* cell growth by inhibiting the DNA gyrase. In a successful Gateway® recombination reaction the lethal *ccdB* gene is replaced by the DNA fragment that is intended to be cloned, so only the vectors that underwent a successful Gateway® recombination reaction may enable *E. coli* to survive in Zeocin selective medium (figure 2.1). Once the *ccdB* gene is replaced by the intended DNA fragment, the vector will be referred to as pENTR™. The amplified DNA fragments were cloned into pCR 2.1-TOPO® (Invitrogen) vectors for storage and manipulation sake.

## 2.5 SOFTWARE AND STATISTICAL METHODS

The sequences of relevant *Arabidopsis* and cassava genes were obtained from Phytozome (<http://phytozome.jgi.doe.gov/pz/portal.html>) and the BLAST analyses comparing the similarity of gene sequences were also carried out with the same database (Goodstein et al., 2012). The data of gene expression levels in *Arabidopsis* in different parts and stages of development were retrieved from the database of eFP browser (Winter et al., 2007). Primers for PCR, qPCR and sequencing were designed with NCBI primer-BLAST (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) and IDT oligo-Analyzer (<https://www.idtdna.com/calc/analyzer>) (Ye et al., 2012). Primers for the verification of T-DNA insert mutants in *Arabidopsis* were designed through T-DNA express: iSect Primers (<http://signal.salk.edu/tdnaprimers.2.html>). The alignments, comparison and phylogenetic tree reconstruction of sequences were carried out using Clustal Omega (McWilliam et al., 2013), Seaview version 4.5.4 (<http://doua.prabi.fr/software/seaview>) (Gouy et al., 2010) and Geneious R8 (Kearse et al., 2012) to exploit each software's unique advantages. Cassava PPD discolouration was quantified and analysed with a MatLab-based programme designed at ETH-Zurich (Vanderschuren et al., 2014). The restriction cleavage sites in relevant plasmids and DNA sequences were searched and verified in NEBcutter V2.0 (<http://nc2.neb.com/nebcutter2>) (Vincze et al., 2003). All quantitative data was analysed and presented with Microsoft Excel 2010. The

significance of data was validated with Student's t-test and ANOVA (analysis of variance) (Box, 1987) and linear regression (Seal, 1967, Tikku, 1971).

## 2.6 PRIMERS

The primers involved in this project and their working details are listed in Appendix I. Primers were synthesised by Invitrogen unless otherwise stated.

## 2.7 BACTERIA AND PLANTS MANIPULATION

### 2.7.1 Antibiotics

The antibiotics and their working conditions are shown in Table 2.2.

Antibiotic	Solvent	Sterilisation	Working concentration
Ampicillin	MilliQ water	Filter sterilisation	100 µg/mL
Chloramphenicol	EtOH	N/A	30 µg/mL
Gentamycin	MilliQ water	Filter sterilisation	25 µg/mL
Hygromycin B (Melford)	MilliQ water	Filter sterilisation	30 µg/mL
Kanamycin	MilliQ water	Filter sterilisation	50 µg/mL (for <i>E. coli</i> )
			100 µg/mL (for <i>A. tumefaciens</i> )
Rifampicin	DMSO	N/A	50 µg/mL
Zeocin (Invivogen)	MilliQ water	Filter sterilisation	50 µg/mL

**Table 2.2 Antibiotics and their working concentrations.**

### 2.7.2 Culture of bacteria

*E. coli* was grown as colonies on LB agar plates (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl and 16 g/L agar, pH = 7.0) with appropriate antibiotics. These plates were incubated at 37°C for 16 hours for the colonies to develop. One single colony was picked with a sterile 10 µL pipet tip and inoculated 5 mL liquid LB medium with appropriate antibiotics. The liquid culture was kept in a shaking incubator at 37°C, 200 rpm for 16 hours for the cells to proliferate.

*A. tumefaciens* colonies were grown on 2YT agar plates (16 g/L tryptone, 10 g/L yeast extract, 5g/L NaCl and 16 g/L agar, pH = 7.0) with appropriate antibiotics. One single colony was picked with a sterile 10 µL pipet tip and used to inoculated 5 mL liquid 2YT medium with appropriate antibiotics. The liquid culture was incubated in a shaking incubator at 28°C, 200 rpm for 40 hours.

### 2.7.3 Preparation of electro-competent *A. tumefaciens*

A single colony of *A. tumefaciens* GV3101 was picked with a sterile 10 µL pipette tip and used to inoculate 5 mL of liquid 2YT medium with appropriate antibiotics. The liquid culture was incubated at 28°C, 200 rpm for 16 hours (overnight incubation was optimal). The next morning 5 mL of the overnight culture was used to inoculate 100 mL of liquid 2YT medium containing appropriate antibiotics. The culture was incubated at 28°C, 200 rpm and the cell density was monitored by reading the OD<sub>600</sub> with a spectrophotometer every 30 mins. When the OD<sub>600</sub> reached 0.9-1.0, the culture was chilled on ice for 30 mins to halt cell growth. The cells were harvested by centrifuging in a pre-chilled Universal 32R centrifuge (Hettich) at 2500 rpm, 4°C for 20 mins. The supernatant was discarded and the cell pellet was washed by re-suspending in 20 mL ice-cold sterile MilliQ water. The cells were pelleted again and the supernatant was discarded. The re-suspension and precipitation were repeated twice and the cells were collected by centrifuging under the same conditions. The cell pellet was re-suspended in 10 mL ice-cold sterile 10% glycerol. The glycerol suspension was pelleted by centrifuging and the supernatant was discarded. The cell pellet was re-suspended in 400 µL of ice-cold sterile 10% glycerol solution. 50 µL of suspended cell was aliquoted to pre-chilled sterile 1.5

mL microfuge tubes and snap-frozen in liquid nitrogen. The cell aliquots were stored at -80°C.

#### **2.7.4 Transformation of chemically competent *E. coli***

Commercial chemically competent cell of 10-beta Competent *E. coli* (High Efficiency) (New England Biolabs) was used. A 50 µL aliquot of competent cell was briefly thawed on ice and mixed with 1-2 µL of suspended vector by gently vortexing of the pipette tip. The aliquot was heat-shocked at 42°C for exactly 30 secs and incubated on ice for 2 mins. The cell aliquot was then mixed with 500 µL of sterile SOC medium and incubated at 37°C, 200 rpm for 60 mins. 100 µL of the culture was spread on an LB agar plate with appropriate antibiotics. The plate was incubated at 37°C for 16 hours for the colonies to develop.

#### **2.7.5 Transformation of electro-competent *A. tumefaciens***

The homemade *A. tumefaciens* electro-competent cells were transformed by the electroporation method. A 50 µL aliquot of *A. tumefaciens* cells was thawed on ice and mixed with 15 ng of purified plasmid DNA. The mixture was immediately transferred to an ice-cold electroporation cuvette (Bio-Rad, 1 mm gap) and electroporated with a MicroPulser (Bio-Rad) with one single pulse at 2.5 kV. 1 mL of sterile 2YT medium was added to the cuvette immediately after the electroporation. The culture was transferred to a sterile 15 mL centrifuge tube and incubated at 28°C, 200 rpm for three hours. 50 µL of the culture was spread on a 2YT agar plate with appropriate antibiotics. The plate was incubated at 28°C for 40 hours for the colonies to develop.

#### **2.7.6 Growth of *Arabidopsis***

Freshly harvested seeds were kept in a desiccator (with silica gel) for one week before using. *Arabidopsis* seeds were surface sterilised by soaking in 70% EtOH for 5 min, then in 50% bleach-0.05% Tween20 solution for 5 min, and washed four

times with sterile distilled water. The treated seeds were suspended in sterile 0.15% agar and kept at 4°C for three days to stratify. The stratified seeds were either sown in soil or on MS-MES agar plates.

*Arabidopsis* seeds were sown on Levington's F2-with-sand compost (treated with 0.02 g/L Intercept 70WG, Scotts). The plants were kept in a controlled environment suite (Weiss-Gallenkamp) at 21°C, 60% relative humidity, and 16 hours of daylight a day. The surface sterile seeds could also be sown on sterile MS-MES plates (4.4 g/L Murashige and Skoog salts with Gamborg's vitamins (Melford), 10 g/L sucrose, 0.5 g/L MES (Melford) and 16 g/L Phytoagar (Duchefa), pH = 5.8). The plates were sealed with Parafilm 'M' (Pechiney) and kept in a controlled environment suite (Weiss-Gallenkamp) at 21°C, 60% relative humidity, and 16 hours of daylight a day. To obtain *Arabidopsis* roots, 100 mm square plates were used and they were kept at 5° from vertical to encourage the growth of roots.

#### **2.7.7 Homozygous T-DNA inserted *Arabidopsis* selection**

The *Arabidopsis* F6'H1 T-DNA insert mutant seeds obtained from NASC were not purely homozygous (with T-DNA inserts in both copies of the F6'H1 gene), and among these seeds heterozygous and wild-type individuals also existed. A screening was required to verify and obtain homozygous individuals.

40 to 80 seeds from each *Arabidopsis* T-DNA insert mutant line were sown on soil and were allowed to grow for two weeks to develop rosette leaves. One single rosette leaf from each individual plant was detached and the DNA was extracted. 1-10 ng of DNA from each individual was used in a genotyping PCR reaction with primer pairs targeting either the wild-type gene or the T-DNA insertion mutation, following the instruction from SALK institute. The individuals showing T-DNA insertion mutation's band but no wild-type band were considered as homozygous and allowed to grow to develop seeds. These homozygous T-DNA insert mutant *Arabidopsis* were used in downstream experiments.

### 2.7.8 *Arabidopsis* transformation and homozygosity selection

Individuals of homozygous *Arabidopsis* T-DNA insert mutants were transformed using an *Agrobacterium*-mediated floral dipping method (Clough and Bent, 1998). Ten individual plants with many unopened flowers (approximately five weeks old) were used in each transformation (named T0). A single colony of *A. tumefaciens* GV3101 carrying a desired expression vector was used to inoculate 5 mL of liquid 2YT medium with appropriate antibiotics and incubated at 28 °C, 200 rpm for 16 hours in a shaking incubator. 2 mL of the culture was added to 100 mL of fresh liquid 2YT medium with appropriate antibiotics and incubated for 16 hours in a shaking incubator at the same conditions. The *Agrobacterium* cells were then harvested by centrifuging in a Universal 32R centrifuge (Hettich) at 2500 x g, room temperature for 20 mins. The cells were re-suspended in transformation buffer (5% sucrose and 0.01M MgCl<sub>2</sub> in sterile distilled water). The cell density in the buffer was optimised by adjusting the OD<sub>600</sub> of the buffer to 1.0. Silwet L-77 surfactant (Lehle Seeds) was added to the buffer to a final concentration of 0.05% (v/v) before floral dipping. The bolts (with flowers) of the plants were immersed into the buffer for 20 secs and shaken in the buffer gently to remove bubbles. The plants were removed from the buffer and were tied to wooden sticks to stand. These transformed plants were covered with a black plastic bag (to create a dark and damp environment) for 23 hours. Then the plants were allowed to grow and develop seeds (named T1). T1 seeds were surface sterilised and mixed with sterile 0.15% agar. The seeds in agar were spread on MS-MES agar plates containing 30 µg/mL hygromycin B for selection (Zhang et al., 2006). These plates were covered with aluminium foil and kept in dark at 24 °C for four days for the seedlings to germinate. Seedlings that developed a long (7-8 mm) hypocotyl were assumed to be transformed successfully. These positive seedlings were transferred to soil and were allowed to grow and develop seeds (named T2).

T2 seeds from each T1 individual were separately collected and desiccated. Approximately 150 T2 seeds from each T1 individual were surface sterilised and sown on MS agar plates containing 30 µg/mL hygromycin B. These plates were kept in dark at 26 °C for four days and the percentage of positive seedlings was analysed. T1 lines showing approximately 75% of positive T2 seedlings were assumed to carry a single insertion. 10-20 positive seedlings from each chosen line were transferred to soil and allowed to develop seeds (named T3). Approximately

100 T3 seeds from each chosen T2 line were surface sterilised and sown on MS agar plates containing 30 µg/mL hygromycin B and kept in dark at 24°C for four days. T3 lines showing 100% positive seedlings were considered to be homozygous. T3 lines with homozygous single T-DNA insertion mutation were used in downstream experiments.

#### **2.7.9 *Arabidopsis* Gus (β-glucuronidase) staining**

The GUS staining method used in this study was modified from Page (Page, 2009). *Arabidopsis* seedlings (grown on MS-MES agar plates) of 1-2 weeks old were washed briefly with distilled water and soaked in staining buffer (100 mM potassium phosphate buffer (pH = 7.0), 1 mg/ml x-gluc, 0.1% Triton X-100, 3 mM potassium ferricyanide, and 3 mM potassium ferrocyanide) at 37°C for one hour. The stained samples were washed with distilled water and de-stained in 70% EtOH until chlorophyll and other pigments were removed. These samples were photographed with a PENTAX K20D camera.

#### **2.7.10 Subculture and growth of cassava**

Both wild-type and transgenic cassava plants were grown from *in vitro*. Buds with 6-10 mm of stem were cut from *in vitro* cassava and subcultured in sterile CBM pots (4.4 g/L Murashige-Skoog medium with vitamins, 20 g/L sucrose, 2 µM CuSO<sub>4</sub>, and 3 g/L Gelrite (Duchefa), pH = 5.9). The subcultures were allowed to grow *in vitro* for 2-4 weeks in a controlled environment suite (24°C, relative humidity 50%, 16 hours daylight, Weiss-Gallenkamp) for the root and stem to develop. The subcultured cassava plantlets were transferred to soil (mixture of three parts of Levington's M2 compost and one part of perlite). These cassava plants were then kept in the University glasshouse with supplemented lighting at 30°C during the day and 17°C during the night.

### 2.7.11 Transformation of cassava

The transformation of cassava was carried out at ETH-Zurich by Dr Ima Zainuddin following the friable embryogenic calli (FEC) transformation method previously developed (Bull et al., 2009). The transformation comprised three steps, namely the generation of FEC, the *A. tumefaciens*-induced transformation, and the regeneration of transgenic cassava plantlets. The regenerated transgenic cassava lines are examined with Southern blotting to verify lines with single-insertion.

## 2.8 NUCLEIC ACID MANIPULATION

### 2.8.1 *Arabidopsis* leaf genomic DNA extraction

The *Arabidopsis* genomic DNA extraction method was modified from Dr Maha Al-jabri (University of Bath, Bath). A small rosette leaf was excised from *Arabidopsis* (1-2 weeks old) and placed in a sterile 1.5 mL microfuge tube with 200  $\mu$ L of DNA extraction buffer (0.14 M d-Sorbitol, 0.22 M Tris-HCl (pH = 8.0), 0.022 M EDTA (pH = 8.0), 0.8 M NaCl, 0.8% CTAB, and 0.1% n-Laurylsarcosine). The leaf was homogenised with a pillar drill and incubated at 65°C for 5 mins to release the DNA. 100  $\mu$ L of chloroform was added to the homogenate and the mixture was vigorously vortexed. The vortexed mixture was centrifuged in a MiniSpin microfuge (Eppendorf) at 13,000 rpm for 5 mins. The aqueous phase was transferred to a fresh 1.5 mL microfuge tube and mixed with 150  $\mu$ L of 100% isopropanol. The mixture was incubated at room temperature for 15 mins and centrifuged at 13,000 rpm for 20 mins to pellet the DNA. The supernatant was discarded and the pellet was washed once with 500  $\mu$ L of 70% EtOH. The pellet was allowed to air-dry for 10 mins and re-suspended in 50  $\mu$ L of MilliQ water. The DNA sample was quality-controlled by reading its  $A_{260}$  and  $A_{280}$  with a spectrophotometer. The concentration of DNA is calculated with the formula below:

$$[\text{DNA}] \text{ (ng/}\mu\text{L)} = A_{260} \times \text{dilution factor} \times 50$$

The  $A_{260}:A_{280}$  ratio indicates the protein contamination in the sample and should be greater than 1.6.



### 2.8.2 *Arabidopsis* RNA extraction

*Arabidopsis* tissue was harvested and ground in a sterile 1.5 mL microfuge tube with a sterile plastic pestle and liquid nitrogen. Total RNA was extracted with a SV Total RNA Isolation System (Promega) following the manufacturer's instruction. RNA was eluted with 100 µL of nuclease free water (Fisher) and the quality was verified with a NanoVue spectrophotometer (GE) by reading the absorbance at 230 nm, 260 nm and 280 nm. The formula below was used to calculate the RNA concentration:

$$[\text{RNA}] \text{ (ng/}\mu\text{L)} = A_{260} \times \text{dilution factor} \times 40$$

The  $A_{260}:A_{280}$  ratio indicates the protein contamination in sample, and ideally should be greater than 1.8. The  $A_{260}:A_{230}$  ratio indicates the chemical carry-over from the extraction buffers such as guanidine or phenol, and ideally should be greater than 2.0.

### 2.8.3 Cassava leaf DNA extraction

200 mg of cassava young leaf was ground to a powder with liquid nitrogen and glass beads in a sterile 1.5 mL microfuge tube and the powder was transferred to a sterile 2 mL microfuge tube. 1 mL of extraction buffer (2% CTAB, 2% PVP, 100 mM Tris-HCl (pH = 8.0), 25 mM EDTA, 2 M NaCl, and 2% β-mercaptoethanol freshly added before use) was added and mixed thoroughly. The homogenate was incubated at 50°C for 15 mins and vortexed every 3-5 mins. The homogenate was centrifuged in a MiniSpin microfuge (Eppendorf) at 13,000 rpm for 5 mins. 900 µL of supernatant was transferred to a fresh 2 mL microfuge tube and mixed with 900 µL of phenol: chloroform (1:1). The mixture was thoroughly vortexed and centrifuged at 13,000 rpm for 10 mins and the aqueous phase was transferred to a fresh 2 mL tube. The aqueous phase was vortexed with 900 µL of chloroform and centrifuged at 13,000 rpm for 10 mins. The aqueous phase was transferred to a fresh 2 mL microfuge tube and the chloroform extraction was repeated once. 650 µL of the aqueous phase was transferred to a fresh 1.5 mL microfuge tube and mixed with equal volume of 100% isopropanol. The solution was incubated at -80°C for 16 hours and centrifuged at 13,000 rpm for 30 mins to pellet nucleic acid. The supernatant was discarded and the pellet was washed once with 70% EtOH. The

pellet was air-dried at room temperature for 10 mins and re-suspended in 100  $\mu$ L MilliQ water. The quality and concentration of the sample were checked with a NanoVue spectrophotometer (GE). The DNA concentration should be approximately 500 ng/ $\mu$ L.

#### **2.8.4 Southern blotting**

The Southern blotting was carried out by Dr Ima Zainuddin (ETH-Zurich, Switzerland). Standard protocols were followed (Sambrook et al., 1989) and consumables from Roche were used. Non-radioactive digoxigenin-labelled probe was prepared with primer pair 2, targeting the hygromycin-resistant gene *hptII*. 10-12  $\mu$ g of DNA from each sample was digested with HindIII (New England Biolabs) and used in the blotting.

#### **2.8.5 Cassava root RNA extraction**

Total RNA was extracted from cassava roots following a method modified from previous literatures (Chang et al., 1993, Yu et al., 2008).

A slice of cassava root sample was chopped to small pieces with a sterile knife and then ground to a fine powder in a sterile mortar and pestle with liquid nitrogen. 100 mg of root powder was transferred to a sterile 1.5 mL microfuge tube and mixed with 600  $\mu$ L of pre-heated RNA extraction buffer (2% CTAB, 2% PVP, 100 mM Tris-HCl (pH = 8), 25 mM EDTA, 2 M NaCl, and 2%  $\beta$ -mercaptoethanol freshly added before use). The mixture was vigorously vortexed and incubated at 65°C for 15 mins, with vortexing every 3-5 mins. The homogenate was chilled on ice and 500  $\mu$ L of phenol: chloroform = 1:1 mixture was added. The mixture was vortexed and centrifuged in a MiniSpin microfuge (Eppendorf) at 13,000 rpm for 15 mins. The aqueous phase was transferred to a fresh 2 mL microfuge tube and vortexed with 500  $\mu$ L of chloroform. The mixture was centrifuged at 13,000 rpm for 10 mins and the aqueous phase was transferred to a fresh 1.5 mL microfuge tube. The aqueous solution was mixed with 1/4 volume (approximately 150  $\mu$ L) of 10M LiCl to make a final concentration of 2 M. The solution was incubated at -20°C for 16 hours and the RNA was pelleted by centrifuging at 13,000 rpm for 20 mins. The

supernatant was discarded and the pellet was re-suspended in 340  $\mu$ L of nuclease-free water (Fisher). 340  $\mu$ L of chloroform was added to the solution and vortexed to remove the carry-over contaminants. The mixture was centrifuged at 13,000 rpm for 10 mins and the aqueous phase was carefully transferred to a fresh 1.5 mL microfuge tube without disrupting the sediments. 34  $\mu$ L of 3 M sodium acetate and 1 mL of EtOH were added and the mixture was vortexed. The solution was incubated at  $-80^{\circ}\text{C}$  for at least four hours, and then centrifuged at 13,000 rpm for 20 mins to pellet the RNA. The supernatant was removed and the pelleted RNA was washed once with 70% EtOH. The pellet was air-dried for 10 mins and re-suspended in 50  $\mu$ L of nuclease-free water (Fisher). 35  $\mu$ L of the re-suspended RNA was digested with Turbo DNase (Ambion) in the following set-up: 35  $\mu$ L RNA, 5  $\mu$ L 5x Turbo DNase buffer, 4  $\mu$ L nuclease-free water and 1  $\mu$ L Turbo DNase (2 U/ $\mu$ L). The reaction was set up in sterile 0.2 mL thin-walled PCR tubes and incubated at  $37^{\circ}\text{C}$  for 30 mins. The enzyme was removed by vortexing with 50  $\mu$ L of chloroform and then centrifuged at 13,000 rpm for 3 mins. The aqueous phase was mixed with 5  $\mu$ L of 3 M sodium acetate and 160  $\mu$ L of EtOH. The solution was incubate at  $-80^{\circ}\text{C}$  for four hours and centrifuged at 13,000 rpm for 20 mins. The pellet was re-suspended in 50  $\mu$ L nuclease-free water and the quality of the RNA samples was verified with a NanoVue spectrophotometer (GE). The RNA samples were stored at  $-80^{\circ}\text{C}$  until needed.

### **2.8.6 *E.coli* plasmid extraction**

5 mL or 50 mL of sterile LB medium with appropriate antibiotics (depending on the amount of plasmid required) was inoculated with a single *E. coli* colony. The culture was incubated in a shaking incubator at  $37^{\circ}\text{C}$ , 200 rpm for 16 hours to allow the cells to proliferate. The cells were harvested by centrifuging in a Universal 32R centrifuge (Hettich) at 2000 x g and the supernatant was discarded after being treated with bleach. Plasmid was extracted from the cell pellet with a QIAprep Spin Miniprep Kit (Qiagen) or a Plasmid Midi Kit (Qiagen) and eluted with appropriate amount of MilliQ water following the instruction from the manufacturer. The quality and concentration of DNA was examined by spectrophotometry as described above.

## **2.8.7 cDNA preparation and quality control**

A High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) with a supplement of oligo dT(18) primer (New England Biolabs) were used to synthesise cDNA from total RNA samples. 150 ng of total RNA was used in each reaction and the instruction from the manufacturer was followed, plus 1  $\mu$ L of 50  $\mu$ M oligo dT (18) primer in each reaction. The reaction was incubated in a Thermo Cyclor PTC-200 (MJ Research).

## **2.8.8 PCR**

### **2.8.8.1 Genotyping/screening PCR**

In this study PCR was widely used as a genotyping and screening tool to determine the presence/absence of a target DNA sequence. Either cDNA or purified genomic DNA was used as template unless otherwise stated.

PCR reactions were set up in 0.2 mL thin-walled PCR tubes. Taq polymerase with standard buffer (New England Biolabs) and dNTP set (Thermo Scientific) were used following the instructions from the manufacturers. Each 25  $\mu$ L reaction contained 1 pg to 10 ng of DNA template (depending on the abundance of the target sequence), 2.5  $\mu$ L 10x standard Taq polymerase buffer, 2  $\mu$ L dNTP mix (2.5 mM each, 10 mM in total), 1  $\mu$ L forward primer (10  $\mu$ M), 1  $\mu$ L reverse primer (10  $\mu$ M), 0.125  $\mu$ L Taq polymerase (5 U/ $\mu$ L) and MilliQ water to make up the volume to 25  $\mu$ L. Reactions were carried out in a Thermo Cyclor PTC-200 (MJ Research) with the following programme:

94 °C for 2 mins,

30x {94 °C for 30 secs,

A °C for 30 secs,

72 °C for B mins},

72 °C for 5 mins.

'A°C' is the annealing temperature which is determined according to the sequences of the primer pair used in this reaction. The 'A°C' of each primer pair is recorded in Appendix I. 'B mins' is the extension time which is determined according to the length of the amplicon. Every kb of base-pairs requires 1 min of extension time.

#### **2.8.8.2 Colony PCR**

Transformed *E. coli* colonies were genotyped with an alternate method of colony PCR to determine the presence/absence of a desired fragment in plasmids they possessed.

700 µL of liquid LB medium with appropriate antibiotics was inoculated with a colony of interest and incubated in a shaking incubator at 37°C, 200 rpm for eight hours. After incubation, 0.5 µL of liquid culture containing suspended cells was added to a genotyping PCR reaction as template. A modified PCR programme was used, as describes below:

94 °C for 5 mins,

30x {94 °C for 30 secs,

A °C for 30 secs,

72 °C for B mins},

72 °C for 5 mins.

An extended denaturing heating was aimed at breaking the cells and thus releasing plasmids. The 'A°C' and 'B mins' are determined as per above.

### 2.8.8.3 Cloning PCR products

When amplifying a DNA fragment for sequencing or vector construction, a high-fidelity polymerase with proof-reading ability was used to minimise errors during amplification.

Q5<sup>®</sup> Hot Start High-Fidelity DNA Polymerase (New England Biolabs) was used when high-fidelity amplification was required. A reaction was set up in a 0.2 mL thin-walled PCR tube and each reaction contained 1 pg to 10 ng of DNA template (depending on the abundance of target sequence), 5 µL 5x Q5<sup>®</sup> polymerase buffer, 2 µL dNTP (2.5 mM each, 10 mM in total), 1 µL forward primer (10 µM), 1 µL reverse primer (10 µM), 0.25 µL Q5<sup>®</sup> polymerase (2 U/µL) and MilliQ water to 25 µL. The reaction was carried out in a Thermo Cycler PTC-200 (MJ Research) with the following programme:

98 °C for 1 mins,

30x {98 °C for 20 secs,

A °C for 20 secs,

72 °C for B mins},

72 °C for 2 mins.

‘A°C’ and ‘B mins’ are determined according to primer-pair combinations as per above.

### 2.8.8.4 ‘Touch-up’ PCR

A ‘touch-up’ PCR programme was used when primers containing additional restriction sites or attB sequences were used in the reactions. Q5<sup>®</sup> Hot Start High-Fidelity DNA Polymerase (New England Biolabs) was used to minimise errors in amplification. The reaction set-up was the same to the cloning PCR and the programme is described below:

98 °C for 1 mins,  
5x {98 °C for 20 secs,  
A °C for 20 secs,  
72 °C for B mins},  
30x {98 °C for 20 secs,  
C °C for 20 secs,  
72 °C for B mins},  
72 °C for 2 mins.

‘A°C’ is the annealing temperature of the primer sequence complementing to the template (‘low’ annealing temperature) and ‘C°C’ is the annealing temperature of the full length of primers (‘high’ annealing temperature). These are determined according to the sequences of primer pairs as shown in Appendix I. ‘B mins’ is the extension time determined as per above.

## **2.9 PLASMID MANIPULATION**

### **2.9.1 DNA restriction endonuclease digestion**

Restriction endonucleases (restriction enzymes) from New England Biolabs were used unless otherwise stated. Up to 1 µg of DNA was used in each digestion unless otherwise stated. Reactions were set up and stopped according to the instruction from the manufacturer. The digested DNA fragments were analysed and purified by agarose gel electrophoresis.

### **2.9.2 DNA electrophoresis and clean-up**

Nucleic acid samples were analysed using agarose electrophoresis. To prepare an agarose gel, 100 mL 1x TAE buffer (40 mM Tris, 20 mM acetic acid, and 1 mM

EDTA) and 0.8 g agarose (Bioline) were mixed and boiled in a microwave oven until the agarose fully dissolved. This agarose solution was cooled to approximately 60°C and mixed with ethidium bromide solution to a final concentration of 0.25 mg/mL. The solution was poured into a sealed gel rack with appropriate combs and allowed to solidify. Solidified gel was immersed in 1x TAE buffer in an electrophoresis tank, and nucleic acid samples containing 1x loading dye (8% sucrose and 0.05% bromophenol blue) were loaded into the wells. The electrophoresis was run at 80 V for 40-100 mins depending on the size of nucleic acid fragments. The gel was visualised with a GDS 7500 UV transilluminator (UVP) and Grab-IT 2.0 imaging software (Synoptics Ltd).

DNA fragments may be extracted and purified from the gel. The DNA band was excised from the gel with a sterile razor and kept in a sterile microfuge tube. A QIAquick Gel Extraction Kit (Qiagen) was used to extract the DNA from the gel following the instruction from the manufacturer. The DNA was eluted in 50 µL of MilliQ water.

### **2.9.3 A-tailing and TOPO<sup>®</sup> TA cloning**

DNA fragments amplified with Q5<sup>®</sup> High-Fidelity DNA Polymerase (New England Biolabs) did not possess the 3' overhanging deoxyadenosine due to the proof-reading ability of the high-fidelity polymerase. To enable TA cloning reaction, A-tailing was required. The PCR product was extracted from agarose gel and 20 µL of purified DNA (concentration varies) was mixed with 2.5 µL 10x Standard Taq polymerase buffer, 2 µL dNTP (2.5 mM each, 10 mM in total), 0.125 µL Taq polymerase (5 U/µL) and MilliQ water to 25 µL in a 0.2 mL thin-walled PCR tube. The reaction was incubated at 72°C for 30 mins and 1 µL of the product was used in TA cloning immediately. TOPO<sup>®</sup> TA Cloning kit (Invitrogen) was used following instruction from the manufacturer. 1 µL of the product was used to transform competent *E. coli* cells immediately.

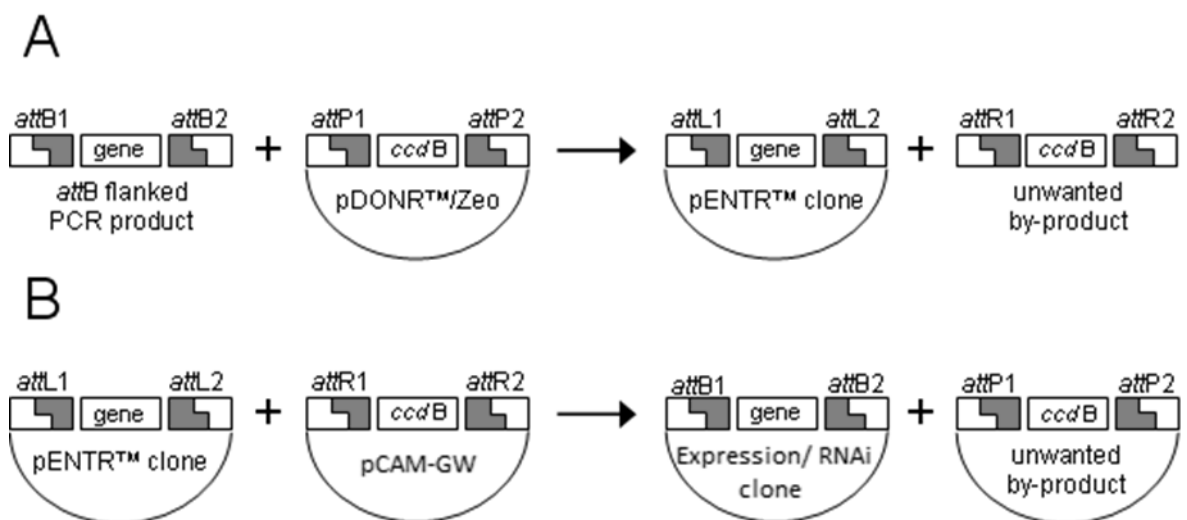


### 2.9.4 DNA ligation

T4 DNA ligase (Thermo Scientific) was used and instructions from the manufacturer were followed with minor modifications. Reactions were set up in 0.2 mL thin-walled PCR tubes and incubated at 16°C for 16 hours. The reaction was terminated by heating at 70°C for 10 mins and 1 µL of the product (containing the resultant vector) was used to transform competent *E. coli* immediately.

### 2.9.5 Gateway® cloning

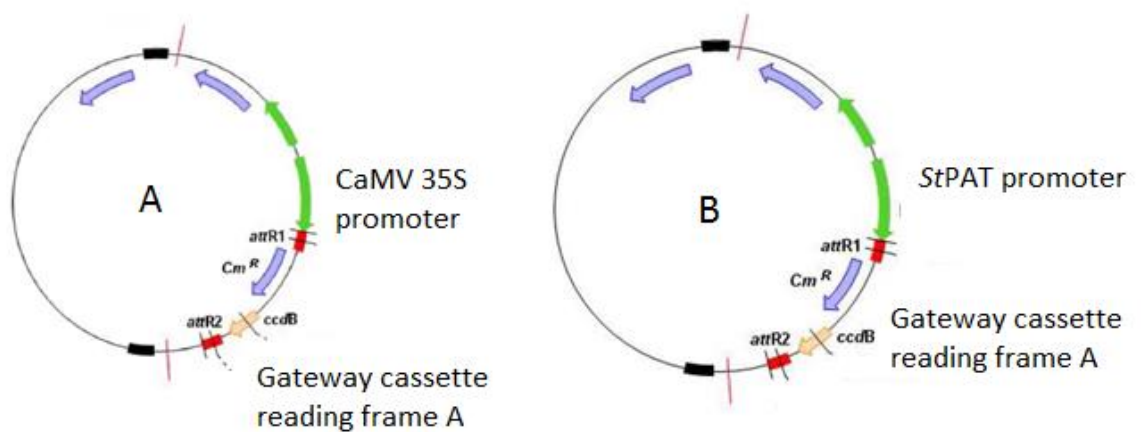
Gateway® cloning technique was used to clone a DNA fragment into a Gateway® competent destination vector without restriction cleavage. The Gateway® reaction is a recombination reaction catalysed by a site-specific recombinase (a lambda integrase family member) to transfer DNA fragments between vectors (Doerks et al., 2002). A series of modified *att* sequences are involved and two recombination reactions carried out by two different recombinases are required for a Gateway® cloning. BP Clonase™ II (Invitrogen) recombines *attB* site with *attP*, while LR Clonase™ II (Invitrogen) recombines *attL* site with *attR* as shown in figure 2.1.



**Fig. 2.1 The method of Gateway® recombinations.** (A) BP Clonase™ II joined the desired fragment to the intermediate vector pDONR/Zeo. (B) LR Clonase™ II joined the desired fragment to the destination vector. GW: Gateway® cassette. Figure modified from the Gateway® Technology with Clonase™ II manual version A, June 2004 (Invitrogen).

The attB sites (full sequences as in primer pair 17) were added to a PCR amplified gene fragment in two 'touch-up' PCR reactions. First reaction was carried out with primers possessing both complementary sequences to the target gene fragments, and 3' half of the attB sequences. The second reaction was with primers of the full length of attB sequence (primer pair 17) and the purified product of the first reaction was used as the template DNA. The attB sites enabled it to be recombined with the attP sites in pDONR/Zeo vector by BP Clonase™ II so that the desired DNA fragment was moved to the pDONR™ vector, converting it to an intermediate named pENTR™. The attP sites were converted to attL sites by the recombination as shown in figure 2.1. The attL sites were recombination targets of LR Clonase™ II which recombined it with the attR sites in the destination vector so that the desired sequence was transferred to the destination vector.

The destination vector was modified from pCAMBIA 1305.1 (with either constitutive CaMV 35S promoter or root-specific *St*PAT promoter) by replacing its GusPlus gene with the Gateway® cassette Reading Frame A (rfA) in the Gateway® Vector Conversion System (Invitrogen) via restriction cleavage and T4 ligation. The resulting Gateway®-competent vectors are shown in figure 2.2. Gateway® competent pCAMBIA vector with *St*PAT was prepared by removing the CaMV 35S promoter and GusPlus through restriction cleavage with enzymes PstI and PmII, and ligating to a PCR amplified *St*PAT fragment and rfA with T4 DNA ligase. Gateway® reactions were carried out with the BP and LR Clonase™ enzyme mixes (Invitrogen) and the instructions from the manufacturer were followed.



**Fig. 2.2 pCAMBIA1305.1 is converted to two kinds of Gateway<sup>®</sup>-competent vectors. (A) CaMV 35S promoter with the Gateway<sup>®</sup> reading frame A cassette; (B) StPAT promoter with the Gateway<sup>®</sup> reading frame A cassette.**

## 2.9.6 DNA sequencing

Plasmids were purified from *E. coli* and amplified PCR products purified from agarose gels as previously described. DNA samples were re-suspended in MilliQ water and dispersed in 1.5 mL microfuge tubes as 15 µL aliquots. Plasmid sample was diluted to 50 ng/µL and PCR product was diluted to 5 ng/µL. The DNA samples and appropriate primers (10 pmol/µL in 15 µL aliquots) were sent to Eurofins MWG Operon to be sequenced.

## 2.10 PPD DISCOLOURATION QUANTIFICATION

### 2.10.1 Cassava root sample preparation

Wild-type and RNAi transgenic cassava plantlets were planted in 9 cm x 9 cm pots and were allowed to grow for 6-8 months in the University glasshouse to develop storage roots. The roots were harvested and cleaned with tap water. The cleaned roots were soaked in 25% bleach for 10 mins and washed three times with sterile distilled water. These roots were then dried with paper tissue and cut into 8-10 mm thick slices. Root slices were kept in Petri dishes with a piece of sterile filter paper (Whatman) lining below. These Petri dishes were closed but not sealed. These root

slices were kept in a controlled environment suite (Weiss-Gallenkamp) at 24°C, relative humidity 50% over a time-course of 5 days to develop discoloration.

### 2.10.2 Quantification of discolouration

Discoloration of cassava slices was quantified with a MatLab-based software developed at ETH-Zurich based on binary histogram (Kunttu et al., 2003, Vanderschuren et al., 2014). The dehydrated surface of the slice was removed with a sterile knife to expose a fresh section, and the slice was photographed against a piece of filter paper (Whatman) with a PENTAX K20D camera (aperture f/5.6, exposure time 1/180 sec and ISO 400) under supplementary lighting. PPD discolouration levels of the parenchyma part of the slice was quantified with the software following the developer's instruction (Vanderschuren et al., 2014). The software converts the colour picture into greyscale and analyses its normalised pixel frequencies to generate a histogram (Zainuddin, personal communication). The grey value that covers 97.5% (97.5% quantile), and grey value that covers 2.5% (2.5% quantile) were used to calculate the score with the formula below:

$$Score = \frac{\text{Quantile 97.5\%} - \text{Quantile 2.5\%}}{\text{Quantile 97.5\%}}$$

The reading from the cassava root sample was divided by the reading from the filter paper beneath, in order to standardise the data. The final score indicated how many folds the reading of the root sample was greater than that of the filter paper, and was calculated with the formula below:

$$Final\ score = \frac{\text{Score sample}}{\text{Score filter paper}}$$

Cassava roots of different discoloration final scores are shown in figure 5.4. Note that, because the score is generated using the filter paper as a standard, a root with no PPD has a score of above zero.

## **2.11 LC-MS BIOCHEMICAL QUANTIFICATION**

### **2.11.1 *Arabidopsis* root sample preparation**

Wild-type and transgenic *Arabidopsis* were allowed to grow for 4 weeks in MS-MES square plates before their roots were harvested. The root was cleaned in MilliQ water and air-dried on a piece of paper towel for one hour. The root sample was weighed and placed in a 1.5 mL microfuge tube and stored at -80°C until needed. The root sample was ground to powder with a plastic pestle and liquid nitrogen before scopoletin extraction.

### **2.11.2 Cassava root sample preparation**

Cassava root samples were cut to small pieces with a sterile knife and ground to fine powder in sterile mortars and pestles with liquid nitrogen. The root powder was stored at -80°C until needed.

### **2.11.3 Scopoletin extraction and LC-MS quantification**

15-30 mg of *Arabidopsis* root powder or 200-700 mg of cassava root powder was transferred to a sterile 2 mL screw-cap microfuge tube (Molecular BioProducts). 2.0 mL of Extraction solvent (HPLC-grade MeOH containing 100 ng/mL 4-MU) was added to the root powder and the tube was incubated at room temperature on a vertical rotor (Labinco B. V.) at 10 rpm for 16 hours. The tube was centrifuged in a Minispin microfuge (Eppendorf) at 10,000 rpm for 5 mins to pellet the root powder. The supernatant was filtered through a Minisart 0.25 µm NY syringe filter (Sartorius) and transferred to a fresh 2 mL screw-cap microfuge tube (Molecular BioProducts). Filtered supernatant was vacuum-concentrated in a Savant SpeedVac vacuum concentrator (Thermo Fisher) for approximately 3.5 hours until the solvent had been completely removed. 200 µL of Re-dissolving solvent (HPLC-grade MeOH containing 1 µg/mL scoparone) was added to the tube and incubated at 4 °C for 48 hours to dissolve the pelleted extract. The re-dissolved sample was vortexed vigorously and centrifuged at 13,000 rpm for 5 mins to pellet any potential particles. 150 µL of supernatant was transferred to a 300 µL screw-cap vial (Waters). The

sample was quantified on a Bruker Daltonics micrOTOF spectrometer using Electrospray Ionisation (ESI) with a reverse phase C18 coated column in the Department of Pharmacology, University of Bath. The mobile phase was gradient MeOH/water as in previous literature (Kai et al., 2006). 4-MU and scoparone are compounds that share similar structure with scopoletin, but are not detected in cassava roots. These two compounds are added to the extraction as internal controls to correct the losses during scopoletin quantification.

## **2.12 QUANTITATIVE RT-PCR AND DATA ANALYSIS**

### **2.12.1 Primer efficiency quality control**

StepOne Software v2.3 was used to design the qRT-PCR programmes, and the amplification efficiencies of primers that were used in qRT-PCR were tested with a standard curve method. Only primer pairs with good quantification ability (R-squared >0.95, i.e. good linearity) and optimal amplification efficiencies (90%-110%) were accepted.

The quality of primer pair's efficiency and specificity was tested before the primers were used. For each primer pair, serially diluted cDNA samples were used to test their qualities. A cDNA sample from fresh wild-type (TMS 60444) cassava root was serially diluted to five concentrations: 1x, 0.2x, 0.04x, 0.008x and 0.0016x. Every primer pair was tested by qRT-PCR with the five serially diluted cDNA samples as templates, and every reaction was performed in triplicate. The amplification efficiency and linearity of quantification were calculated by the StepOne software. The amplification efficiency was worked out with the formula below:

$$E\% = (10^{(-1/\text{slope})} - 1) \times 100\%$$

The slope was worked out by plotting the quantitative data points from the five serial dilutions of cDNA templates to draw a standard curve. A primer pair was accepted if the trendline of the data points fitted to an R-squared greater than 0.95, and amplification efficiency between 90% and 110%.

The quantitative PCR reactions were set up on ice in a MicroAmp Fast Optical 96-Well Reaction Plate (Applied Biosystems) as following: 1 µL cDNA (from approximately 7.5 ng cassava total RNA), 10 µL Fast SYBR Green mastermix

(Applied Biosystems), 0.4 µL forward primer (10 µM), 0.4 µL reverse primer (10 µM) and 8.2 µL nuclease-free water (Fisher) for each reaction. The plate was sealed with MicroAmp Optical Adhesive Film (Applied Biosystems) following the manufacturer's instructions, and centrifuged at 1,000 rpm for 1 min to eliminate bubbles. The qRT-PCR was carried out in a StepOnePlus™ quantitative PCR machine (Applied Biosystems) with the programme below:

95 °C for 20 secs,

40x {95 °C for 3 secs,

60 °C for 30 secs},

95 °C for 15 secs,

60 °C for 60 secs,

And then the temperature was increasing 0.3°C every 15 secs until 95°C to create a melting curve.

### 2.12.2 Quantitative PCR

The cDNA samples were standardised with the quantitative data of the cassava house-keeping gene, Polyubiquitin-10 (GenBank accession number DV441403) with primer pair 11, as the reference gene (Czechowski et al., 2005, Dekkers et al., 2012).

The reactions were set up as that in the primer quality control section. A cDNA sample prepared from fresh wild-type (TMS 60444) cassava root was used as the reference sample. The target genes were quantified using the  $\Delta\Delta C_t$  method (Livak and Schmittgen, 2001). The  $C_t$  values of each reaction were collected and the relative quantity (RQ) of gene expression in each sample was calculated referring to the reference sample with the formula below:

$$\text{Relative quantity of target gene} = 2^{-(C_t(\text{Target gene}) - C_t(\text{Reference gene}) - \Delta C_t(\text{Reference sample}))}$$

In which:

$$\Delta C_t (\text{Reference sample}) = C_t (\text{Target gene of reference sample}) - C_t (\text{Reference gene of reference sample})$$

# 3 DO F6'H1 CANDIDATE GENES EXIST IN CASSAVA GENOME?

## 3.1 INTRODUCTION

### 3.1.1 Scopoletin and PPD

In this study our focus is to find cassava genes that regulate the 6'-hydroxylation of feruloyl CoA, a crucial step in the biosynthesis of scopoletin. We believe that scopoletin plays an important role in the post-harvest physiological deterioration (PPD) of cassava; therefore this study ultimately addresses the mechanism of PPD.

After harvesting, various fluorescent hydroxycoumarins, including scopoletin, scopolin, esculetin, and esculin, accumulate in the parenchyma tissue of cassava roots. Wheatley first discovered that after harvest, the scopoletin amount in cassava root increased 150 to 200-fold in 24-48 hours, which strongly indicates that scopoletin is involved in PPD development (Wheatley, 1982). This result was confirmed by others, showing that the dramatic accumulation of metabolites consists mainly of scopoletin and scopolin (Buschmann et al., 2000b). As part of this process, the activity of phenylalanine ammonia lyase (PAL), the key entry enzyme for phenylalanine into general phenylpropanoid metabolism that includes scopoletin *de novo* synthesis, also increases during PPD response (Gomez-Vasquez et al., 2004, Owiti et al., 2011). Another indication of scopoletin's role in PPD is that both the fluorescence emission and the blue-black discoloration initiate around the xylem vessels at the early stages of PPD (Buschmann et al., 2000b). Scopoletin forms a blue-black precipitation *in vitro* when mixed with hydrogen peroxide and peroxidase, thereby confirming its major role in the discoloration observed during the PPD (Beeching et al., 2002, Gutiérrez-Mellado et al., 1996). Further supporting evidence for the oxidation of scopoletin contributing to the PPD is that exclusion of oxygen through waxing the roots effectively delays PPD development (Reilly et al., 2003).

Applying scopoletin to fresh cassava roots can significantly accelerate the vascular and parenchymal discoloration (Wheatley and Schwabe, 1985), which may be due to a boost of oxidizing substrate, or due to a potential signalling role of scopoletin.



However, since there is no definite evidence supporting scopoletin's signalling function in plants, the latter interpretation is weak. Therefore, scopoletin probably does act as a quencher to the ROS components such as  $H_2O_2$ , which starts accumulating during the wound-triggered oxidative burst (Reilly et al., 2003). After a week of storage, scopoletin levels decline in cassava roots, which may be due to that the cassava roots are too damaged or exhausted to replenish scopoletin consumed in PPD response (Buschmann et al., 2000b).

### 3.1.2 Scopoletin biosynthesis in cassava and *Arabidopsis*

Three alternate pathways for the biosynthesis of scopoletin within phenylpropanoid metabolism are known in plants (figure 1.11, pathways 1, 2, and 3). All the three pathways require phenylalanine ammonia-lyase (PAL) to convert phenylalanine, which is a primary metabolite, to cinnamate. To simplify the explanation, coenzyme A is omitted from the metabolites. Cinnamate is converted to p-coumarate, from where the three possible pathways diverge. From p-coumarate to scopoletin, five reactions are required: 3'-hydroxylation, 3'-methylation, 6'-hydroxylation, trans/cis isomerisation of 1'-double bond, and lactonisation. Every pathway undergoes all the five steps, but in different orders. In pathway (1), p-coumarate is first 6'-hydroxylated (to form 2', 4'-dihydroxycinnamate), and then E-Z isomerised ((Z)-2',4'-dihydroxycinnamate), lactonised (umbelliferone), 3'-hydroxylated (esculetin), and finally 3'-methylated to form scopoletin. In pathway (2) and (3), p-coumarate is first 3'-hydroxylated to caffeate. Then in pathway (2), caffeate is further 6'-hydroxylated (6'-hydroxycaffeate), isomerised ((Z)-6'-hydroxycaffeate), lactonised (esculin), and finally 3'-methylated to scopoletin; whilst in pathway (3), caffeate is 3'-methylated (ferulate), 6'-hydroxylated (6'-hydroxyferulate), isomerised ((Z)-6'-hydroxyferulate), and finally lactonised to form scopoletin. By using isotope-labelled intermediates in the pathways, it has been proved that pathway (3) synthesises more than 90% of the scopoletin present in cassava roots (Bayoumi et al., 2010, Bayoumi et al., 2008a). During the growth of cassava roots, scopoletin may be converted to scopolin, its biochemically inert  $\beta$ -D-glucopyranoside, by covalently linking a glucose molecule on its hydroxyl group by UDP-Glc:phenylpropanoid glucosyltransferase, to be stabilised and solubilised (Ahn et al., 2010, Bayoumi et al., 2008b, Chong et al., 2002).

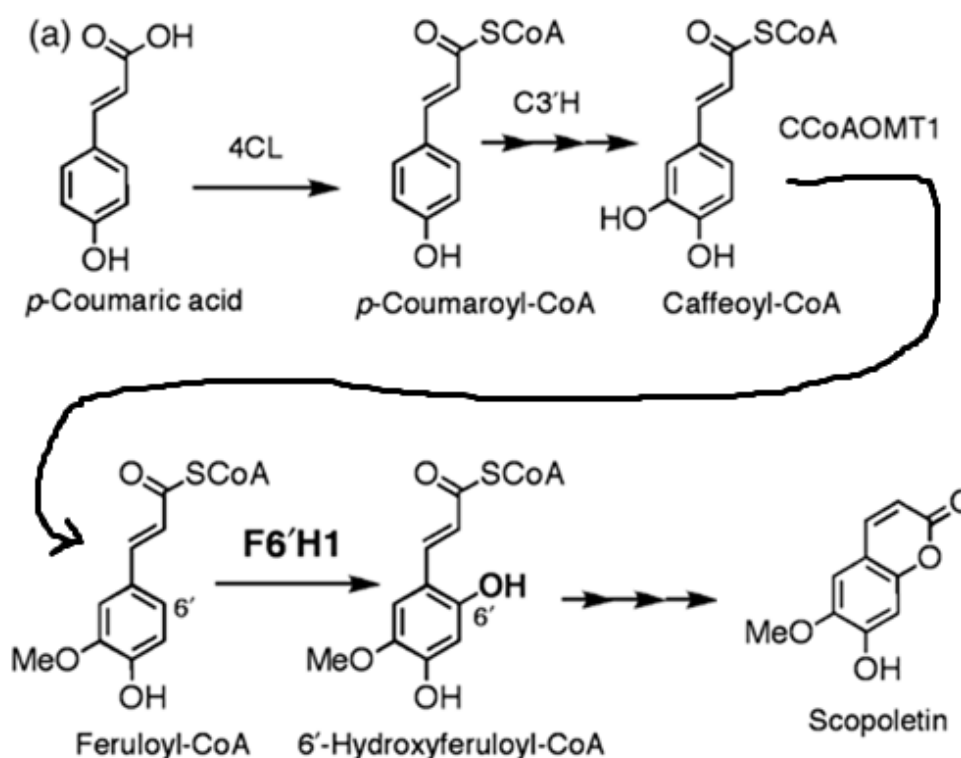
The accumulated scopoletin in cassava roots possibly comes from two sources: the *de novo* biosynthesis from phenylpropanoid metabolism, and the decomposition of scopolin, the glycoside of scopoletin. Scopolin is biologically inactive (as a phytoanticipin) as its hydroxyl group is protected by the glycosidic bond, so that the plant is able to keep a 'pool' of scopoletin ready to use without disrupting the normal metabolism. Evidence indicates that scopoletin may be stored and when required,  $\beta$ -glucosidase can hydrolyse scopolin to release scopoletin and contribute to the anti-stress response (Ahn et al., 2010, Gutiérrez-Mellado et al., 1996). However, the exact roles of scopoletin and scopolin in PPD response are yet unclear, and it is an interesting question to be addressed in this study.

### **3.1.3 *Arabidopsis* scopoletin biosynthesis and F6'H1**

Although the roles of possible alternative pathways in scopoletin biosynthesis have not been rigorously explored in *Arabidopsis*, it is probable that it shares the same major routes as cassava and tobacco, as shown in figure 3.1 (Bayoumi et al., 2008a, Bayoumi et al., 2008b, Fritig et al., 1970, Kai et al., 2006). Since phenylpropanoid metabolism emerged early in the evolution of terrestrial plants to enable them to adapt to the novel challenges and opportunities of life out of water (Tohge et al., 2013), the majority of its diverse pathways are conserved amongst higher plants. Like cassava, *Arabidopsis* also synthesises scopoletin and scopolin in its roots, which gives us a convenient tool to study the genes regulating scopoletin synthesis in this model plant (Kai et al., 2006). The *Arabidopsis* gene F6'H1 (feruloyl CoA 6'-hydroxylase 1, Phytozome gene ID: At3g13610, NCBI Reference Sequence: NC\_003074.8), which regulates the 6'-hydroxylation of feruloyl CoA, is required for scopoletin synthesis in *Arabidopsis*, as when it is knocked out through T-DNA insertion mutation, scopoletin cannot accumulate in *Arabidopsis* roots (Kai et al., 2008). These data confirm that the predominant scopoletin biosynthetic pathway in *Arabidopsis* is the same as pathway (3) in cassava (Kai et al., 2008).

F6'H1 belongs to the 2-oxoglutarate dependent dioxygenase (2OGD) enzyme family. In the presence of Fe (II) ion, 2OGD incorporates both atoms of an O<sub>2</sub> molecule to a 2-oxoglutarate molecule and its substrate. In the case of scopoletin

biosynthesis, F6'H1 introduces a hydroxyl group to feruloyl-CoA, converting it to 6'-hydroxyferuloyl-CoA.



**Fig. 3.1 A part of the *Arabidopsis* scopoletin biosynthetic metabolism.** The probable pathway, from *p*-coumarate to scopoletin, is shown in this picture with abbreviation of major enzymes of each step. 4CL: 4-coumarate coenzyme A ligase; C3'H: caffeate 3'-hydroxylase; CCoAOMT1: caffeoyl CoA O-methyltransferase; F6'H1: feruloyl CoA 6'-hydroxylase 1. Picture modified from Kai et al. (2008).

### 3.1.4 Aim of research

This discovery in *Arabidopsis*, that a T-DNA insertion mutation in the F6'H1 gene is sufficient to prevent scopoletin synthesis in the roots, suggests a strategy through which the role of scopoletin in cassava PPD response can be studied. The major pathway of scopoletin biosynthesis is shared between the two plants. Therefore, a candidate gene (or a candidate gene family) of *Arabidopsis* At3g13610 (F6'H1) which controls the 6'-hydroxylation of feruloyl CoA, must exist in cassava. If this candidate gene or genes were identified in the cassava genome, RNAi constructs

may be prepared with the potential of inhibiting the expression of one or more of those candidate genes in cassava. Therefore, the research in this chapter is focussed on identifying those cassava candidates of *Arabidopsis* F6'H1.

## 3.2 RESULTS

### 3.2.1 Cassava has a small family of F6'H1 candidate genes

The *Arabidopsis* gene encoding F6'H1 (At3g13610) has 1169 bp from start codon to stop codon. The gene harbours one 83 bp intron between two similar-sized exons (512 bp and 574 bp respectively). The predicted cDNA and amino acids sequences of F6'H1 were retrieved from the database of Phytozome (Goodstein et al., 2012). The sequences were used to search in the genome of cassava (also within the database of Phytozome) for homologous sequences. BLASTN and TBLASTN methods were used for the cDNA and amino acid sequences respectively with the default algorithmic parameters (Altschul et al., 1990, Gertz et al., 2006). The hits (identified similar gene sequences) were evaluated by two parameters, the score and the *E*-value. The score indicates how similar the hit is to the query sequence, and the *E*-value indicates the possibility that the hit is due to the enormous size of the database (Goodstein et al., 2012). Ideally a reliable hit should have high score and low *E*-value. Although there is no clear cutoff between 'good' and 'bad' hits, we chose the hits with scores greater than 200 and *E*-values less than  $1\text{E-}50$  ( $1 \times 10^{-50}$ ) for further analysis. The top hits retrieved from the cDNA and amino acid sequence searches are shown in table 3.1 and table 3.2.

These data indicate that regardless of which sequence (either cDNA or amino acid) is used as the query sequence, the top hits are the same group of genes. The homogenous search identified seven gene sequences in the cassava genome: cassava4.1\_010376m.g, cassava4.1\_010292m.g, cassava4.1\_010291m.g, cassava4.1\_033240m.g, cassava4.1\_010381m.g, cassava4.1\_027567m.g, and cassava4.1\_030526m.g. The remaining hits after these seven have much lower scores and much higher *E*-values, indicating that the seven genes above are all the cassava genes with the highest similarity to *Arabidopsis* F6'H1. These seven genes are referred to as 10376, 10292, 10291, 33240, 10381, 27567, and 30526 respectively, for simplicities sake. The gene or genes that regulate the

6'-hydroxylation of feruloyl CoA in cassava are likely to be within these seven genes, but which ones were the functional F6'H1 orthologs was unclear. To have a clearer image, the similarity between *Arabidopsis* F6'H1 sequence and each individual cassava gene is compared. The score of similarity between these genes and F6'H1 was calculated with Clustal Omega (McWilliam et al., 2013) and the result is shown in table 3.3 and table 3.4. It can be concluded from the results that the cassava genes were more similar to each other than to the *Arabidopsis* F6'H1. This makes it difficult to decide which of cassava genes were the most related one to the F6'H1. To gain a further insight into the relationship between *Arabidopsis* F6'H1 and the cassava candidate genes, we constructed phylogenetic trees.

cDNA alignment			
Sequence names	Scores	E-values	Locus names
<b>Scaffold09151</b>	358.4	7.50E-97	cassava4.1_027567m.g
			cassava4.1_010381m.g
<b>Scaffold08673</b>	349.3	3.90E-94	cassava4.1_010376m.g
			cassava4.1_010292m.g
			cassava4.1_010291m.g
			cassava4.1_033240m.g
<b>Scaffold03476</b>	215.9	5.80E-54	cassava4.1_030526m.g

**Table 3.1 F6'H1 candidate genes are identified from the cassava genome.**

Enquiry with cDNA sequences provides a group of seven genes, as shown in the 'locus names' column.

Amino acid alignment			
Sequence names	Scores	E-values	Locus names
<b>Scaffold08673</b>	502.3	7.10E-159	cassava4.1_010376m.g
			cassava4.1_010292m.g
			cassava4.1_010291m.g
			cassava4.1_033240m.g
<b>Scaffold09151</b>	493.8	8.40E-156	cassava4.1_010381m.g
			cassava4.1_027567m.g
<b>Scaffold03476</b>	278.1	8.20E-129	cassava4.1_030526m.g

**Table 3.2 F6'H1 candidate genes are identified from the cassava genome.**

Enquiry with amino acid sequences provides a group of seven genes, as shown in the 'locus names' column.

Genes	F6'H1	33240	10292	10291	10381	27567	30526	10376
<b>F6'H1</b>	100	67.42	67.42	67.98	67.7	67.13	64.43	63.94
<b>33240</b>	67.42	100	98.34	98.61	92.8	92.52	86.98	86.91
<b>10292</b>	67.42	98.34	100	99.17	92.24	92.56	86.43	86.63
<b>10291</b>	67.98	98.61	99.17	100	92.52	92.84	86.98	86.91
<b>10381</b>	67.7	92.8	92.24	92.52	100	98.34	86.43	86.63
<b>27567</b>	67.13	92.52	92.56	92.84	98.34	100	85.6	86.07
<b>30526</b>	64.43	86.98	86.43	86.98	86.43	85.6	100	91.14
<b>10376</b>	63.94	86.91	86.63	86.91	86.63	86.07	91.14	100

**Table 3.3 The percentage similarity scores between F6'H1 and cassava candidates amino acid sequences.** The numbers shown are the similarities to each other in percentage.

<b>Genes</b>	<b>F6'H1</b>	<b>33240</b>	<b>10292</b>	<b>10291</b>	<b>10381</b>	<b>27567</b>	<b>30526</b>	<b>10376</b>
<b>F6'H1</b>	100	66.67	66.48	66.29	66.76	66.95	66.48	66.1
<b>33240</b>	66.67	100	97.24	97.61	88.67	88.95	82.14	82.5
<b>10292</b>	66.48	97.24	100	97.8	88.58	89.01	82.41	82.41
<b>10291</b>	66.29	97.61	97.8	100	88.21	88.64	81.95	82.31
<b>10381</b>	66.76	88.67	88.58	88.21	100	97.51	82.5	82.78
<b>27567</b>	66.95	88.95	89.01	88.64	97.51	100	81.95	82.59
<b>30526</b>	66.48	82.14	82.41	81.95	82.5	81.95	100	86.65
<b>10376</b>	66.1	82.5	82.41	82.31	82.78	82.59	86.65	100

**Table 3.4 The percentage similarity scores between F6'H1 and cassava candidates cDNA sequences.** The numbers shown are the similarities to each other in percentage.

### **3.2.2 At least four of the candidate genes were expressed in cassava roots during PPD**

It is possible that the cassava F6'H1 candidates family of seven genes contains pseudogenes. A pseudogene is a genomic DNA sequence which shares sequence similarity with its functional counterparts, but is itself functionally inactive (Poliseno et al., 2010). Pseudogenes are created in gene duplication or retrotransposition events. In fact, pseudogenes are so common in the genome that they are likely to be found in most gene families. Since they are non-functional, pseudogenes in cassava F6'H1 candidate family could be excluded from the study. Although none of the seven cassava candidates carry characteristic sequences of a pseudogene, such as polyadenine tail or genetic lesions (e.g. in-frame termination codon or frameshift mutation), confirmation of their expressions is still essential (Vanin, 1985). Non-expressed, probable pseudogenes can be identified via RT-PCR, which differentiates between transcribed and non-transcribed genes.

The predicted coding sequences (i.e. cDNA) of the seven cassava candidates were retrieved from the database of Phytozome. Gene-specific primers were designed based on the specific sequences of each gene, which should be able to differentiate the genes, and cDNAs derived from them, from each other. RNA from cassava roots on 0 day, 1 day and 3 days after harvest was used to prepare cDNA, and primer pairs 33 through 38 were used for the genotyping RT-PCR. To facilitate primers annealing to their correct targets, the PCR protocol was modified. The concentration of primers was lowered to 0.1  $\mu$ M (one fourth the concentration of that in a routine genotyping PCR) in the PCR reaction to reduce non-specific annealing. The RT-PCR was run at a series of annealing temperatures to evaluate the different behaviours between reactions using either genomic DNA or cDNA as templates. The results are shown in figure 3.2 and table 3.5.

If a primer pair amplifies a clear band from the genomic DNA template but not from the cDNA template at a certain annealing temperature, it implies no expression of the gene. If the bands of both genomic and cDNA of a gene disappear at the same temperature as the annealing temperature increases, it is likely that the cDNA band is amplified specifically from the cDNA of this gene, as the primer pair shows the same affinity to genomic DNA and cDNA (figure 3.2 A). However, in some cases the bands were faint regardless of the annealing temperature used, and thus these genes failed to provide convincing evidence. These genes were considered as 'unclear results' (figure 3.2 A). If the genomic DNA bands disappeared before those from the cDNA as annealing temperature increases (figure 3.2 B), although it implies a firm binding between the primer pair and the cDNA, the suspicious behaviour renders these situations 'probable expression'. If the genomic and cDNA bands continue to be detected at high annealing temperatures (figure 3.2 C), this implies a 'strong evidence of expression'.

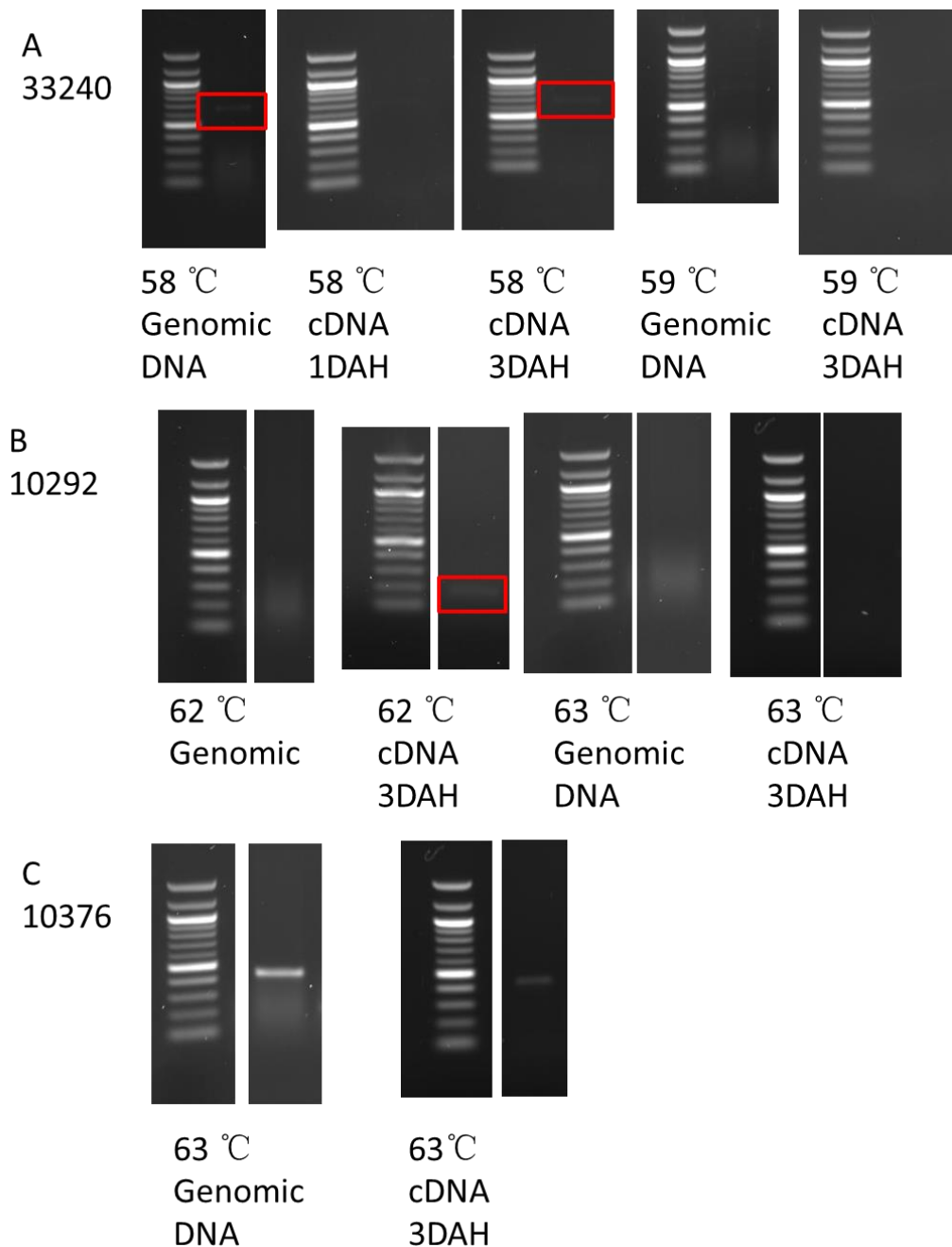
Figure 3.3 and 3.4 showed very high similarity between the seven cassava candidates. The DNA sequences of these genes were very similar, but not identical, and it was extremely difficult to design gene specific primers targeting only one single gene. The primers we designed to differentiate between these genes were often differed by only one single nucleotide. Therefore, primers designed for one gene may also be able to amplify other genes that are very similar to the target one, which makes the unambiguous interpretation of the gene expression results difficult; hence the modification to the PCR protocol. In the RT-PCR results, most genes did



not provide a clear-cut 'yes-or-no' answer to their expression status. There was clear evidence showing genes 10376 and 27567 were expressed during PPD, as their amplification from cDNA always behaved the same as that from genomic DNA. Genes 10291 and 10292 are probably expressed, but the results are less clear. There were no strong and clear bands indicating the expression of gene 10381, 30526 and 33240. Interestingly, none of cassava candidates showed strong expression in freshly harvested roots (0 DAH). This result indicated that these cassava feruloyl CoA 6'-hydroxylase related genes were PPD inducible genes that were activated after the wounding and PPD signalling.

<b>Genes</b>	<b>0 DAH</b>	<b>1 DAH</b>	<b>3 DAH</b>
<b>33240</b>	0	1	1
<b>10381</b>	0	1	1
<b>30526</b>	1	1	1
<b>10376</b>	0	3	3
<b>10292</b>	0	2	2
<b>10291</b>	0	2	2
<b>27567</b>	0	3	3

**Table 3.5 The expression of the seven cassava F6'H1 candidates in different stages of PPD process.** Numbers indicate the likelihood of expression: 3, strong evidence of expression; 2, probable expression; 1, unclear results (very faint or inconsistent bands); 0, not expressed. DAH: days after harvest.



**Fig. 3.2 The presence of bands at different annealing temperatures revealed information on gene expression.** Temperatures: the annealing temperatures of reactions; Genomic DNA or cDNA: the templates used in PCR; DAH (days after harvest): the stage of sample being used in RNA extraction. (A) Bands of gene 33240 appear in genomic DNA and cDNA 3DAH at 58°C, and both disappear at 59°C; Bands never appear if cDNA 1DAH is used as template. (B) Bands of cDNA of gene 10292 appear at 62°C and disappear at 63°C, but bands of genomic DNA start blurring at 62°C. (C) Bands of gene 10376 are clear at high temperature (63°C). Very pale bands are indicated with red boxes.

### 3.2.3 The phylogenetic tree revealed the relation between the cassava candidates gene family and F6'H1

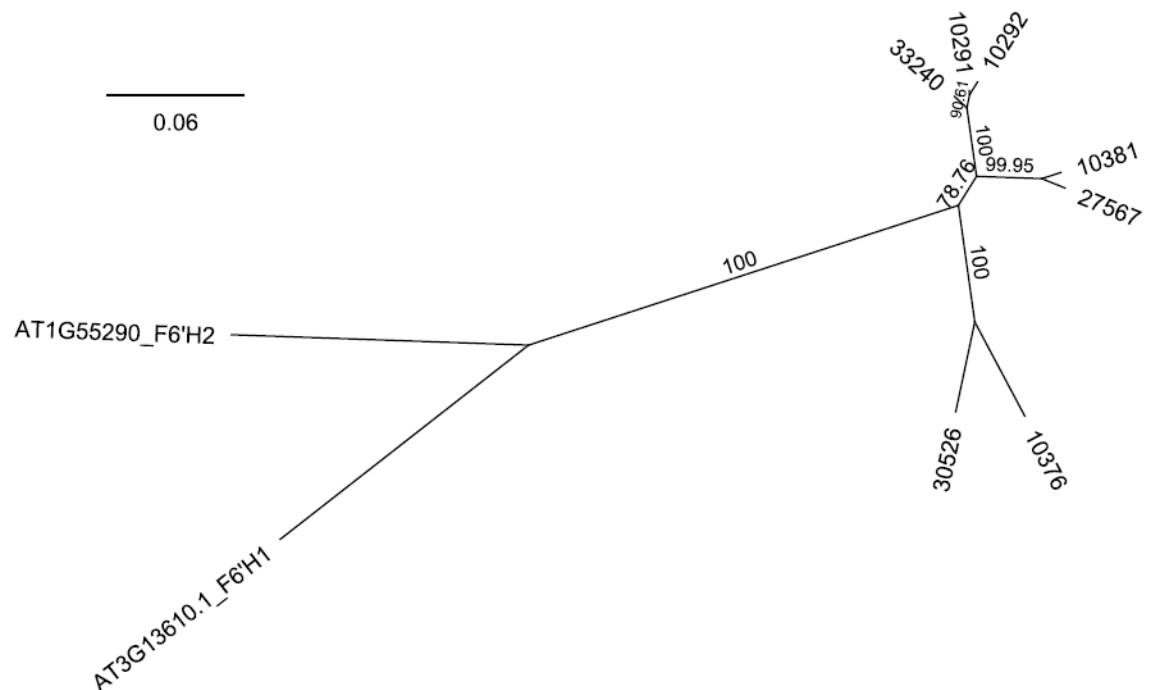
To gain a deeper insight into the relation between *Arabidopsis* F6'H1 and its cassava candidate gene family, the amino acid sequences of six top hit *Arabidopsis* homologues of F6'H1 (AT1G55290, AT3G12900, AT1G04380, AT3G21420, AT5G12270, and AT4G22880), were retrieved from the *Arabidopsis* genome (in Phytozome database) using F6'H1 as the query sequence, by a similar process as described above for the cassava candidates. AT1G55290 encodes another feruloyl CoA 6'-hydroxylase (named F6'H2) which is closely related to F6'H1 but with limited expression strength (Kai et al., 2006). A comparison between F6'H1 and its homologues in *Arabidopsis* and cassava genomes helped us observe the relation between F6'H1 and its orthologs in cassava. In this study we are interested in the function of the cassava genes instead of the DNA sequence, therefore to minimise the impact from neutral mutations, amino acid sequences were used for the tree reconstruction.

The sequence alignment and phylogenetic trees were constructed using Geneious R8 (Kearse et al., 2012). The software uses a neighbour joining (NJ) method to calculate the relation between different sequences (Saitou and Nei, 1987). Bootstrapping was repeated 10,000 times and other parameters were kept as default. Unrooted trees were used to minimise the visual bias in data presentation. The reconstructed phylogenetic tree is shown in figure 3.3. This phylogenetic tree (figure 3.3) confirms that the seven cassava candidate genes are more closely related to each other than to the *Arabidopsis* sequences. Additionally, the two *Arabidopsis* feruloyl CoA 6'-hydroxylases (F6'H1 and F6'H2) are most similar to each other and more similar to the cassava candidates than to the other *Arabidopsis* hits.

An examination of the similarity between the two *Arabidopsis* F6'Hs and the seven cassava candidates (figure 3.4) further confirms these relationships. However, although it is difficult to draw categorical interpretations based on this tree, cassava gene 10376 and 30526 do appear to be closer to F6'H1 than the other candidates, albeit the differences in distances are marginal. Certainly, these trees do not provide sufficient evidence to enable one to conclude definitely that one of these cassava candidates is more likely to have feruloyl CoA 6'-hydroxylase activity over the others.



candidate gene with any degree of certainty, genes 30526 and 10376 were chosen as the cassava candidates that were the most similar to F6'H1.



**Fig. 3.4 Phylogenetic tree (with bootstrapping) showing the similarities between *Arabidopsis* F6'H1 and cassava candidate genes.** This tree is based on amino acid sequences.

### 3.2.4 The genes to be used in downstream experiments were selected

Putting the gene expression and phylogenetic data together, among the cassava candidates, the clearly expressed genes are 10291, 10292, 10376 and 27567, and the genes most similar to F6'H1 are 10376 and 30526. The genes to complement scopoletin biosynthesis in *Arabidopsis*, and the genes to knock out in cassava, are to be chosen from genes 10291, 10292, 10376, 27567 and 30526. To complement scopoletin synthesis in *Arabidopsis* T-DNA insert mutants, constructs based on the entire coding sequences of the cassava genes were necessary. For designing an effective RNAi construct, a deeper look into the sequences of these genes was required in order to identify a sequence that not only was capable of effective down-regulation of the gene from which it was designed, but also with the potential of doing so to other related members of the cassava gene family.

Despite the ambiguous results in gene expression and phylogenetic tree construction, gene 10376 was selected for this purpose due to its higher similarity to *Arabidopsis* F6'H1 as shown in figure 3.3 and 3.4 (and thus more likely to be the F6'H1 homolog controlling feruloyl CoA 6'-hydroxylation), and due to its clear evidence of expression. DNA sequence alignments of 10376 were aligned with the remaining six genes and it revealed long stretches of identical sequences. The genomic DNA of 10376 consisted of two exons of approximately 500 bp each, separated by a small intron of 129 bp, which structure was similar to that of *Arabidopsis* F6'H1. The first and second exons of the cassava gene were aligned and compared (McWilliam et al., 2013) as shown in table 3.6 and table 3.7, which showed that overall the second (3') exon was slightly more conserved than the first (5'). However, in the first (5') exon there are more small stretches of conservative nucleotides in rows compared to the second (3') exon in which the conservative nucleotides are 'allocated' to shorter fragments as shown in figure 3.5. For this reason the first exon of gene 10376 was selected as the basis for the RNAi construct, as it might be able to knock out more than one of the related genes through targeting a shared highly conserved sequence, as evidence showing external RNA may also silence genes with partial complementarity to it (Jackson et al., 2003, Ossowski et al., 2008).

Genes 10291, 10292 and 10376 were chosen to complement *Arabidopsis* F6'H1 T-DNA insertion mutations. The aim was to study if more than one cassava candidates in the gene family were able to complement the function of F6'H1 and the decision was made before the revision of the gene expression data in which gene 27567 was also proved to be expressed. For this reason, gene 27567 was not used in the complementation of scopoletin biosynthesis.

<b>5' part</b>	<b>10291</b>	<b>33240</b>	<b>10292</b>	<b>10381</b>	<b>27567</b>	<b>10376</b>	<b>30526</b>
<b>10291</b>	100	98.22	98.42	88.95	88.56	81.46	82.05
<b>33240</b>	98.22	100	98.62	89.74	89.55	81.46	82.05
<b>10292</b>	98.42	98.62	100	89.74	89.35	81.95	82.45
<b>10381</b>	88.95	89.74	89.74	100	96.84	80.24	81.07
<b>27567</b>	88.56	89.55	89.35	96.84	100	79.76	79.88
<b>10376</b>	81.46	81.46	81.95	80.24	79.76	100	85.82
<b>30526</b>	82.05	82.05	82.45	81.07	79.88	85.82	100

**Table 3.6 A comparison between the sequence similarity of the first exon (5' half) of the seven cassava F6'H1 candidates.** The first (5') exon and the second (3') exon of these genes are compared separately. The numbers in the tables indicate the percentage of similarity between two gene sequences.

<b>3' part</b>	<b>10381</b>	<b>27567</b>	<b>33240</b>	<b>10292</b>	<b>10291</b>	<b>30526</b>	<b>10376</b>
<b>10381</b>	100	98.1	87.74	87.56	87.56	83.42	84.46
<b>27567</b>	98.1	100	88.43	88.72	88.72	83.59	84.19
<b>33240</b>	87.74	88.43	100	96.03	97.06	81.69	83.94
<b>10292</b>	87.56	88.72	96.03	100	97.26	82.21	82.99
<b>10291</b>	87.56	88.72	97.06	97.26	100	81.69	83.33
<b>30526</b>	83.42	83.59	81.69	82.21	81.69	100	86.87
<b>10376</b>	84.46	84.19	83.94	82.99	83.33	86.87	100

**Table 3.7 A comparison between the sequence similarity of the second exon (3' half) of the seven cassava F6'H1 candidates.** The first (5') exon and the second (3') exon of these genes are compared separately. The numbers in the tables indicate the percentage of similarity between two gene sequences.





### 3.3 DISCUSSION

The goal of this chapter was to identify the candidate gene or genes to *Arabidopsis* F6'H1 in cassava genome from which to produce an RNAi construct for the purpose of inhibiting scopoletin biosynthesis in cassava roots post-harvest. BLAST search of the cassava genome using *Arabidopsis* F6'H1 identified seven strong candidate genes with highly similar sequences. Phylogenetic trees derived from their amino acid sequences show that all seven are approximately equally similar to *Arabidopsis* F6'H1 gene and fail to identify any one as the strongest candidate, though genes 10376 and 30526 appear the most attractive. Determination of the expression profiles of the seven candidates via RT-PCR revealed that four genes (10291, 10292, 10376, and 27567) showed strong expression during PPD; therefore, the remaining three could be discarded. While it would have been interesting to test all the candidate genes further, for logistical purposes it was decided to test the function of genes 10376, 10292, and 10291 for their ability to complement F6'H1 T-DNA insertion mutation of *Arabidopsis* and base the production of an RNAi construct to inhibit scopoletin biosynthesis in cassava on 10376. Gene 27567 was not included due to the reason discussed in section 3.2.2. These experiments are described in the following chapters.

# **4 DO CASSAVA F6'H1 CANDIDATES COMPLEMENT SCOPOLETIN BIOSYNTHESIS IN *ARABIDOPSIS* F6'H1 T-DNA INSERTION MUTANTS?**

## **4.1 INTRODUCTION**

### **4.1.1 Background**

In chapter 3, we identified a small family of cassava genes in cassava genome that may be orthologs of *Arabidopsis* F6'H1. Although their sequence similarity suggests that these cassava genes are strong candidates to carry out the feruloyl CoA 6'-hydroxylation activity, it is not definite evidence. Kai et al. used *Arabidopsis* F6'H1 T-DNA insert mutants to confirm the necessity of gene F6'H1 in the biosynthesis of scopoletin (Kai et al., 2006). Therefore, the same mutant may also be used to test the functional identity of these cassava candidate genes through their ability to complement the F6'H1 mutation in *Arabidopsis*, thereby restoring scopoletin biosynthesis. The goal of this chapter is to use this approach to confirm whether members of the cassava F6'H1 gene family possess feruloyl CoA 6'-hydroxylation activity or not.

### **4.1.2 Strategy**

To achieve this goal the following strategies were adopted:

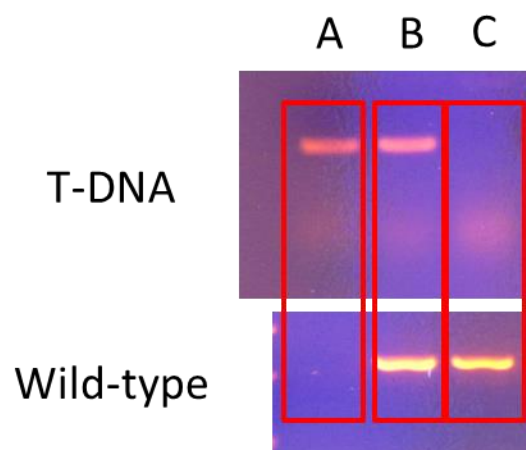
- 1, Confirmation of the *Arabidopsis* F6'H1 T-DNA insertion mutants;
- 2, Cloning cassava candidates of F6'H1 into a T-DNA vector driven by an appropriate promoter;
- 3, Transformation of the T-DNA constructs into *Arabidopsis* mutant lines and confirmation of the transgenesis;
- 4, Determination of scopoletin levels in wild-type and cassava candidate-inserted T-DNA mutant *Arabidopsis* lines;

## 5, Confirmation of the functional identity of cassava candidate genes.

### 4.2 RESULTS

#### 4.2.1 Confirmation of the *Arabidopsis* F6'H1 T-DNA insert mutants

Seeds of three lines of *Arabidopsis* F6'H1 T-DNA insert mutants, namely, SALK\_129938, SALK\_005232, and SAIL\_1252\_A10, were obtained from the Nottingham *Arabidopsis* stock centre (NASC). To achieve the planned scopoletin biosynthetic complementation it was necessary to obtain homozygous mutants for the T-DNA insert lines, so that both copies of F6'H1 genes were inactivated. Therefore, 40 seeds from each line were sown on soil and their genomic DNA was genotyped by PCR using primer pair 5, 4 and 6 targeting the wild-type allele of lines Salk\_129938, Sail\_1252\_A10, and Salk\_005232, whilst primer pair 7, 8, and 9 targeting the T-DNA inserted allele of lines Salk\_129938, Sail\_1252\_A10, and Salk\_005232 respectively. With these primer pairs, their amplification products enabled us to identify homozygous wild-type, heterozygous T-DNA mutants and homozygous T-DNA mutants, as shown in figure 4.1.



**Fig. 4.1 Genotyping PCR on genomic DNA extracted from several *Sail\_1252\_A10* individuals revealed their genotypes.** (A) a homozygous mutant with T-DNA inserts in both copies of F6'H1 gene; (B) a heterozygous mutant with one copy of wild-type F6'H1 and one copy of T-DNA insert mutant; (C) a wild-type plant with two copies of wild-type F6'H1 genes.

The genotyping PCR revealed that there were no homozygous T-DNA insert mutant individual in any line, and that Salk\_005232 only contained wild-type individuals. A replacement was obtained from NASC, but the wild-type problem persisted, so line Salk\_005232 was excluded from the following experiments. Heterozygous mutant individuals of Salk\_129938 and Sail\_1252\_A10 were allowed to self-fertilise and develop seeds. Seeds from these plants were sown again and genotyped as described above so that homozygous F6'H1 T-DNA insert mutants were identified and obtained from both Salk\_129938 and Sail\_1252\_A10.

To verify the inactivation of the F6'H1 gene in these homologous F6'H1 T-DNA insert mutants, cDNA from the roots of Salk\_129938 and Sail\_1252\_A10 seedlings were used as template in RT-PCR. Primer pair 4, which amplifies the 3' of the wild-type gene F6'H1, was used to detect the expression of F6'H1 gene. The results showed that in Sail\_1252\_A10, expression of F6'H1 was not detectable. However, amplification was detected in Salk\_129938, which could be due to the location of the T-DNA insert not preventing transcription (data not shown). To test this hypothesis, primer pair 10, which targets the F6'H1 gene carrying the T-DNA insert, was used in RT-PCR to test for the presence of the T-DNA insert. The result indicated that in line Salk\_129938, the F6'H1 gene was transcribed with the T-DNA inserted in the 3' part (data not shown), which agrees with the location of this T-DNA insert in the database of SALK institute (Alonso et al., 2003). The mRNA containing a large T-DNA insert should not be translatable into a functioning enzyme. Therefore, these data confirmed the identity of the homozygous lines of Salk\_129938 and Sail\_1252\_A10, which were used for subsequent transformation and scopoletin quantification.

#### **4.2.2 Cloning cassava cDNA of F6'H1 candidates into T-DNA vectors driven by suitable promoters**

To express the cassava F6'H1 candidates in F6'H1 T-DNA inserted *Arabidopsis*, the appropriate vector and promoters are required. The vector should contain *Agrobacterium* recognisable terminal repeats and selectable markers (Hoekema et al., 1983). These allow the cassava F6'H1 candidate genes (referred to as the target genes) to be integrated into the host genome, and the successfully transformed host plants to be identified. The target genes also need to be regulated

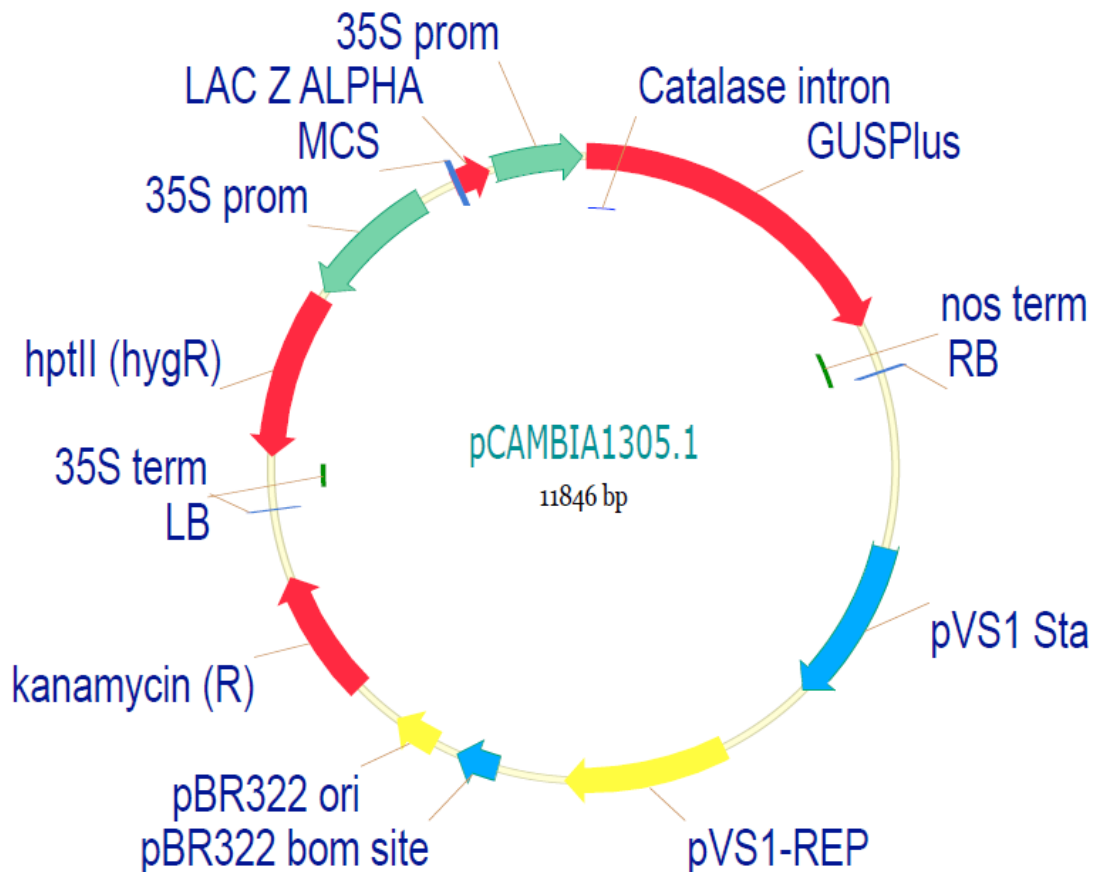
by an appropriate promoter to be expressed in desired parts of the transformant. To facilitate the selection of successfully transformed plants, a hygromycin resistant gene (as a selectable marker) was also necessary in the *Agrobacterium*-compatible vector. The target genes were to be regulated with either a constitutive promoter to guarantee expression throughout the plant, or a root-specific promoter to direct the gene to the cassava roots where it was required. An easy and universal method to clone the target genes into a vector was also required.

Gateway<sup>®</sup> cloning technique was adopted as a universal method to clone the cassava candidate genes (and also the RNAi construct in chapter 5) into the desired expression vector. Gateway<sup>®</sup> cloning technique involves site-specific recombination which enables cloning a DNA fragment into a Gateway<sup>®</sup>-competent vector effectively without complicated restriction site designing (Kalidas et al., 2011).

pCAMBIA 1305.1 plasmid (NCBI Genbank accession number: AF354045.1) is a vector designed by Cambia, Australia which meets our requirements. The structure of pCAMBIA 1305.1 is shown in figure 4.2. Between the two terminal repeats (LB and RB), the vector includes a hygromycin-resistant gene (hptII) as a selectable marker and a GusPlus reporter gene, both under the control of constitutive cauliflower mosaic virus 35S promoters (CaMV 35S). There are convenient restriction cleavage sites before and after the promoter and the reporter gene. This vector can be converted to a Gateway<sup>®</sup>-competent vector in a few simple steps. The reporter gene and/or the promoter was removed via restriction cleavage, to confer the appropriate promoter and a Gateway<sup>®</sup> cassette so that the target genes may be cloned into the vector by Gateway<sup>®</sup> cloning (Karimi et al., 2002).

Two different promoters were chosen to regulate the expression of the target genes. The CaMV 35S promoter, which shows constitutive expression and high activity in both *Arabidopsis* and cassava, is from the cauliflower mosaic virus 35S transcript (Vanderschuren et al., 2007a, Wilkinson et al., 1997). This promoter enables a gene to be constitutively expressed in all parts of a plant (Kay et al., 1987). The CaMV 35S promoter could help us obtain a reliable expression of the cassava genes in transformed *Arabidopsis*. *StPAT* is the promoter regulating the expression of *Solanum tuberosum* storage protein *patatin* in potato tubers (Jefferson et al., 1990). This promoter has been proved to promote gene expression in *Arabidopsis*

roots, and is predominantly root-specific (Page, 2009). Since *Arabidopsis* F6'H1 was found to be mainly expressed in roots (database of eFP browser) (Toufighi et al., 2005, Winter et al., 2007), a root-specific promoter may help us look into the function of the cassava candidates in the *Arabidopsis* roots.



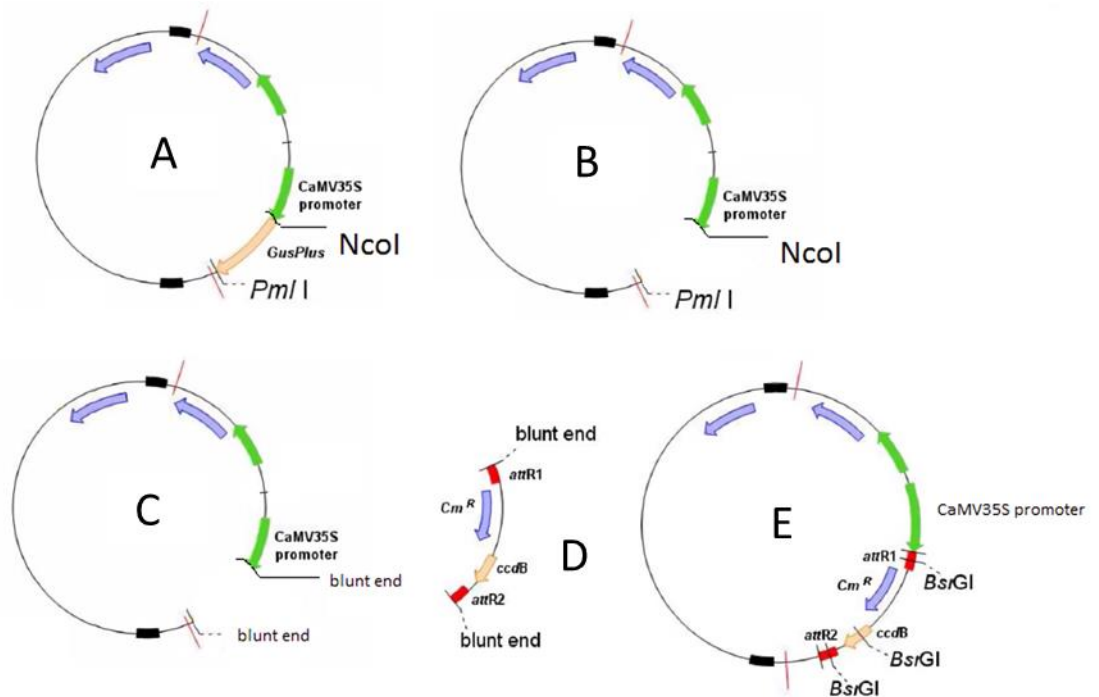
**Fig. 4.2 A map of the binary vector pCambia 1305.1.** LB: left border (terminal) repeat; RB: right border (terminal) repeat. The fragment between LB and RB are inserted into the host genome. hptII: hygromycin-resistant gene hygromycin phosphotransferase; GUSPlus:  $\beta$ -glucuronidase reporter gene. Kanamycin  $\epsilon$ : kanamycin-resistant gene. Picture modified from <http://www.cambia.org/daisy/cambia/585.html>

To clone the target genes into the pCambia 1305.1 vector, and to control these cassava F6'H1 candidates with either a constitutive CaMV 35S or a root-specific StPAT promoter, pCambia 1305.1 vector was modified. As shown in figure 4.3, the pCambia 1305.1 vector was digested with restriction enzyme NcoI and PmlI to

release the GusPlus reporter gene (figure 4.3 A and B). The linearised vector fragment (approximately 9,800 bp) was separated from the GusPlus fragment (approximately 2,046 bp) and purified by agarose gel electrophoresis. The linearised vector was recovered from the gel and the 5' overhang was filled in with DNA polymerase I Klenow fragment (New England Biolabs) to create blunt ends (figure 4.3 C).

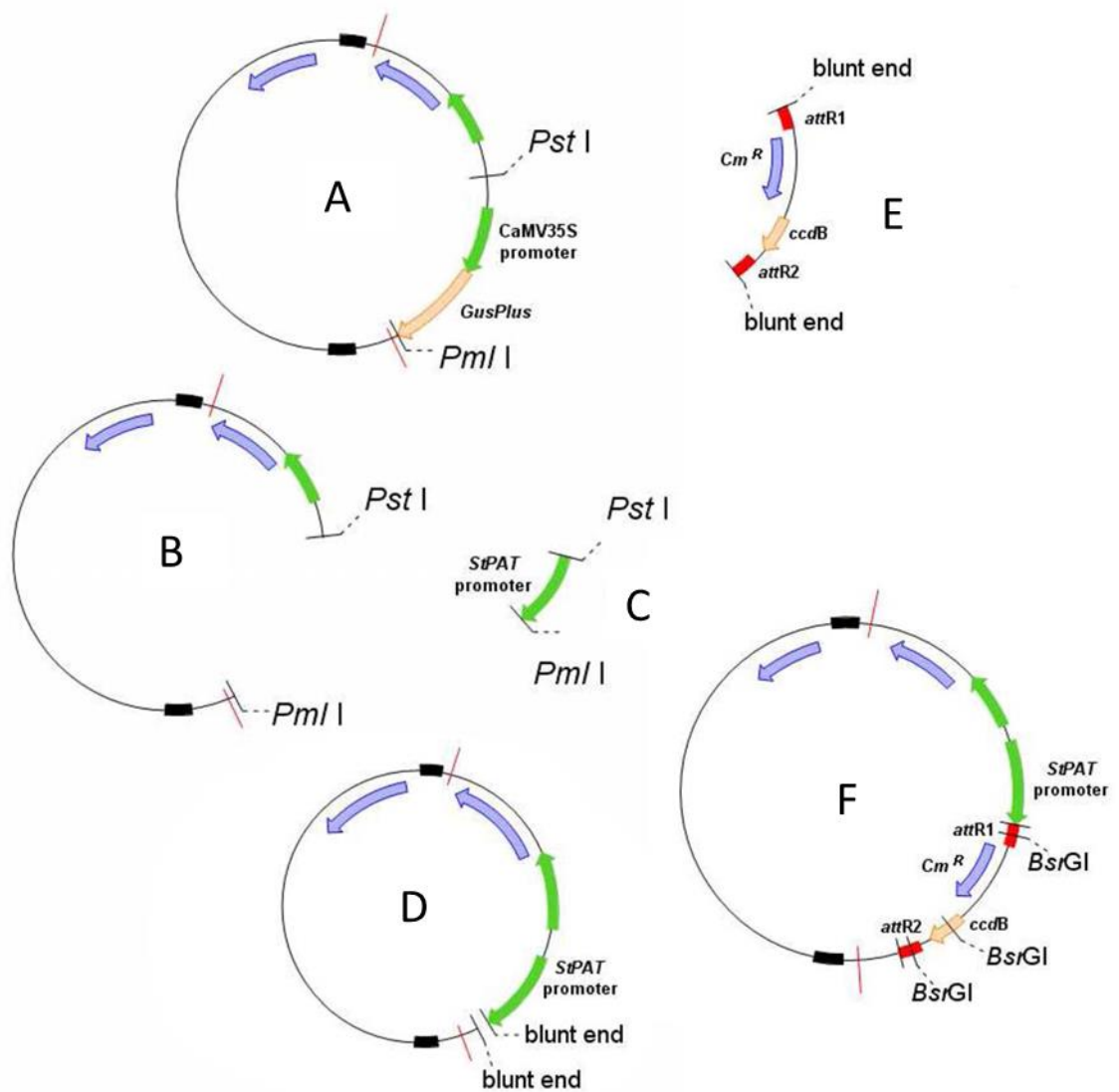
The blunt linearised vector was mixed with the reading frame A (rfA) of the Gateway<sup>®</sup> cassette and ligated (figure 4.3 C, D and E). The product was used to transform One Shot<sup>®</sup> ccdB Survival<sup>™</sup> 2 T1R Competent Cells (Invitrogen<sup>™</sup>), which can survive the presence of the lethal gene ccdB in the Gateway<sup>®</sup> cassette. The cells were grown on LB agar plates with kanamycin and chloramphenicol which were the selectable markers for pCAMBIA 1305.1 and the Gateway<sup>®</sup> cassette, respectively. Surviving colonies were used in a colony PCR with primer pair 25 (which amplifies from the middle of CaMV 35S promoter to the beginning of the Gateway<sup>®</sup> cassette) to confirm the presence and orientation of the Gateway<sup>®</sup> cassette insertion. A correctly constructed vector should produce an amplicon of approximately 323 bp. The vectors that passed this colony PCR test were partially sequenced with primer >35s-rfa-Fwd and >35s-rfa-Rvs to confirm the correctness of the sequence. This resulting vector was termed pCAM:35S:GW.

Gateway<sup>®</sup> competent pCAMBIA 1305.1 vector with the StPAT promoter instead of CaMV 35S was constructed by Bull (Bull, 2011) as shown in figure 4.4. The pCAMBIA 1305.1 vector was digested with restriction enzymes PstI (which cuts once before the CaMV 35S promoter) and PmlI (which cuts right after the GusPlus gene) to remove the promoter and the reporter gene (figure 4.4 A and B). The product was electrophoresed on agarose gel to separate the linearised vector (approximately 9,028 bp) and the promoter-reporter fragment (approximately 2,818 bp). The linearised vector was purified from the gel and ligated with a PCR amplified StPAT fragment (designed to carry the sticky ends of PstI and PmlI) (figure 4.4 B, C, and D). The intermediate vector carrying StPAT promoter was digested with PmlI to linearise the vector right after the StPAT promoter (figure 4.4 D). The Gateway<sup>®</sup> cassette rfA fragment was ligated with the linearised vector so that the Gateway<sup>®</sup> cassette was under the control of the StPAT promoter (figure 4.4 D, E and F). This vector was termed as pCAM:PAT:GW.



**Fig. 4.3 Conversion of pCAMBIA1305.1 vector to a Gateway® competent vector.** (A) the pCAMBIA 1305.1 vector is cleaved with *Nco*I and *Pml*I to remove the *GusPlus* gene; (B) the linearised pCAMBIA 1305.1 vector without the *GusPlus* gene; (C) the sticky ends of the vector is filled in to form blunt ends; (D) the Gateway® reading frame A (rfA) with *attR* sequences and the *ccdB* lethal gene; (E) linearised pCAMBIA 1305.1 vector and the Gateway® reading frame A are ligated. Picture modified from Bull, 2011.





**Fig. 4.4 The pCambia1305.1 vector was converted to Gateway<sup>®</sup> competent vector with StPAT promoter.** (A) the pCambia 1305.1 vector is cleaved with *Pst*I and *Pml*I to remove the *GusPlus* gene; (B) the linearised pCambia 1305.1 vector without the CaMV 35S promoter and the *GusPlus* gene; (C) the *StPAT* promoter with the stick ends of *Pst*I and *Pml*I; (D) the linearised pCambia 1305.1 vector and the *StPAT* fragment are ligated and cut again with *Pml*I to create a blunt end after the *StPAT* promoter; (E) the Gateway<sup>®</sup> reading frame A (rfA) with *attR* sequences and the *ccdB* lethal gene; (F) linearised pCambia 1305.1-*StPAT* vector and the Gateway<sup>®</sup> reading frame A are ligated. Picture modified from Bull, 2011.

As discussed above, three cassava F6'H1 candidates (genes 10291, 10292, and 10376) were considered active genes when the decision was made as to which to test via scopoletin biosynthesis complementation of F6'H1 T-DNA insert mutants of *Arabidopsis*. Gene 27567 was not tested. The selected cassava F6'H1 candidate genes were PCR amplified and attB sequences were added to their ends by 'touch-up' PCR amplification with Q5<sup>®</sup> high-fidelity polymerase (New England Biolabs). Gene 10376 was amplified with primer pairs 13 and 17; gene 10292 was with primer pairs 14, 15, and 17 in a nested PCR; and gene 10291 was amplified with primer pairs 16 and 17. The amplified genes with attB sites were cloned into pCR 2.1-TOPO<sup>®</sup> vector (Invitrogen) to facilitate manipulation. These amplified fragments were electrophoresed and purified from agarose gel, and then A-tailed with Taq polymerase (New England Biolabs) to facilitate TA cloning into the pCR 2.1-TOPO<sup>®</sup> vector followed by transformation into *E. coli*. The cloned cassava genes were then sequenced with primers >M13-fwd and >M13-rvs (M13 universal primers that anneal upstream and downstream of the target sequences in the pCR 2.1-TOPO<sup>®</sup> vector) to verify that the sequence of cassava genes and attB sites were correct.

To perform a Gateway<sup>®</sup> cloning, these pCR 2.1-TOPO<sup>®</sup> vectors were linearised with HindIII, which cuts once in the pCR 2.1-TOPO<sup>®</sup> vector but does not cut within the cassava genes. The linearised vectors were used in the Gateway<sup>®</sup> BP reaction with the Gateway<sup>®</sup> pDONR/Zeo vector and Gateway<sup>®</sup> BP clonase II following manufacturer's instructions. The product was used to transform 10-beta competent *E. coli* (New England Biolabs) and the *E. coli* was grown on LB agar plates with Zeocin. The surviving colonies, which were referred to as the pENTR vectors, should contain the correct gene construct in their vectors.

The pENTR vectors containing the target genes were used in the Gateway<sup>®</sup> LR reaction catalysed by Gateway<sup>®</sup> LR clonase II. Each cassava gene was recombined with both pCAM:35S:GW and pCAM:PAT:GW, creating expression vectors that were controlled by either a constitutive CaMV 35S promoter or a root-specific StPAT promoter. The vectors were sequenced with primer >Attb-FWD and >Attb-RVS (full length attB sequences) to verify the correctness of the sequence of the cassava genes. The vectors carrying correct cassava genes sequence were then used in the transformation of the *Arabidopsis* F6'H1 T-DNA insert mutants.

#### 4.2.3 *Arabidopsis* T-DNA insert lines were transformed with cassava F6'H1 candidate genes

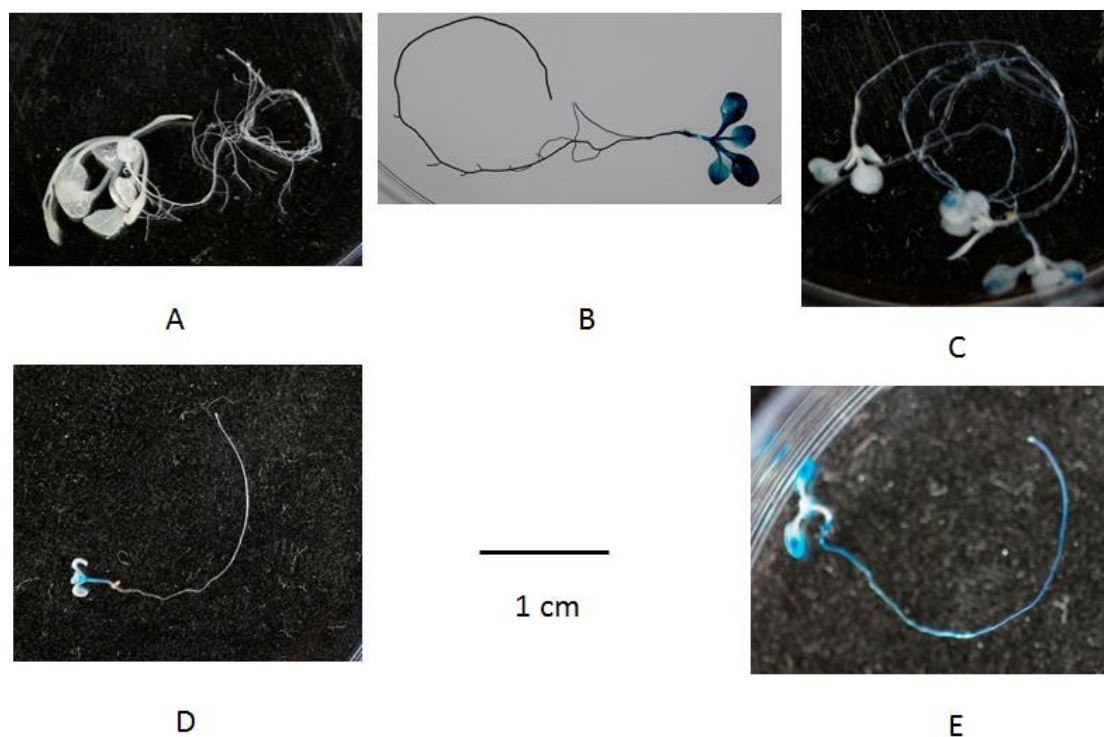
The pCAM:35S:cassava-candidate and pCAM:PAT:cassava-candidate vectors were used to transform *Agrobacterium* GV3101 by electroporation, and these transformed *Agrobacterium* carrying different promoters and target genes were used to transform *Arabidopsis* lines, namely Salk\_129938 and Sail\_1252\_A10, with homozygous T-DNA inserts in the F6'H1 gene. pCAMBIA 1305.1 vector with reporter gene GusPlus and either a *St*PAT or a CaMV 35S promoter were also used in *Arabidopsis* transformation for subsequent use as controls.

All resulting seeds (T1) were screened on MS-MES plates with hygromycin (figure 4.5). Resistant seedlings were transferred to soil and allowed to develop seeds (T2). T2 seeds were again screened on MS-MES plates with hygromycin to verify the segregation of their parents. Lines showing 75% hygromycin-resistant rate were selected and approximately 10 positive individuals from each line were allowed to bear seeds (T3) for hygromycin screening. Lines showed 100% hygromycin-resistant rate were identified and taken to be homozygous for both the F6'H1 T-DNA insert mutant and the cassava transgene.



**Fig. 4.5 Wild-type and hygromycin-resistant *Arabidopsis* seedlings behave differently in a hygromycin screening.** The positive (upper-right) and negative (lower-left) seedlings in a hygromycin screening are indicated by red arrowheads.

To check the success of the transformation and that the two promoters worked in *Arabidopsis* background, the *Arabidopsis* lines transformed with GusPlus reporter gene under the control of either CaMV 35S or StPAT were Gus-stained to assess the activity of the transgene. Seeds with homozygously inserted GusPlus gene were sown on MS plates and allowed to grow for approximately 10 days. The seedlings were stained with Gus stain buffer (figure 4.6). The CaMV 35S promoter showed strong activity in some lines (e.g. figure 4.6 B) but weak activity in others (e.g. figure 4.6 C), which is probably due to the location of inserts in the host genome. In higher plants different regions of the chromosome exhibit different levels of expression activity, which may affect the expression of the transgene (Matzke and Matzke, 1998). However, the activity of the StPAT promoter-regulated lines was not as expected. Most of these did not show root-specific staining. The majority of the transformants did not show detectable Gus activity, while a few showed Gus activity in unexpected parts of the plant, such as the hypocotyl or the tips of rosette leaves (data not shown). As a promoter, StPAT is activated by wounding, stresses and sucrose, in parts where it does not usually exhibit activity (Page, 2009). The *Arabidopsis* seedlings were grown on agar plates, rather than soil, an environment which could be stressful to the plant; additionally, the medium contained high levels of sucrose, which may contribute to the unexpected behaviour of this promoter. On the other hand, unexpected expression of StPAT regulated genes in transgenic plants has also been detected by others (Rocha-Sosa et al., 1989). Since we did not generate many transformants with StPAT promoter and GusPlus gene, our small number of lines may not be representative of the possible range of expression levels.

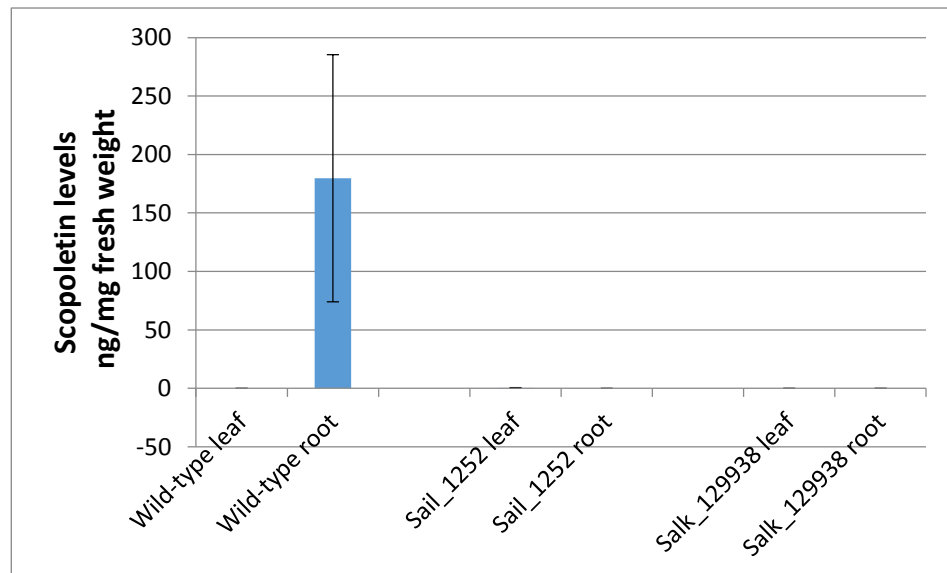


**Fig. 4.6 Gus staining revealed the location of the expression of the GusPlus transgene.** (A) untransformed Sail\_1252\_A10 (as negative control) does not show any staining; (B) CaMV 35S-Gus line 5 shows constitutive activity throughout the plant; (C) CaMV 35S-Gus line 9L shows limited activity in parts of the plant; (D) StPAT-Gus line 1 shows limited activity in the hypocotyl but not the root; (E) StPAT-Gus line 7 shows activity throughout the plant.

#### 4.2.4 Determination of scopoletin levels in wild-type, T-DNA insert mutants *Arabidopsis* lines

Wild-type, Salk\_129938, and Sail\_1252\_A10 *Arabidopsis* seeds were sown on MS-MES plates and were allowed to grow for 4 weeks. Scopoletin levels in the roots and leaves of these plants were quantified by LC-MS (figure 4.7). Scopoletin was present in wild-type roots but not leaves, and was not detected in either the roots or the leaves of the T-DNA insert mutants. However, trace amounts of scopoletin (approximately 0.25 ng/mg FW) were detected in a few samples of Sail\_1252\_A10 leaves, which might be due to slight contamination during sample preparation in these assays. Otherwise it might be explained by the phenylpropanoid metabolic enzymes using accumulated intermediate metabolites from the blocked scopoletin biosynthetic pathway (Bell-Lelong et al., 1997,

Cochrane et al., 2004, Schoch et al., 2001, Toufighi et al., 2005, Winter et al., 2007). For example, Kai et al. also found a second F6'H1-like gene named F6'H2 (Phytozome ID: AT1G55290) which exhibits a background expression in the leaves like many other phenylpropanoid metabolic enzymes (Kai et al., 2008).



**Fig. 4.7 Scopoletin accumulated in wild-type *Arabidopsis* roots but not F6'H1 T-DNA insert mutants.** Data are mean levels of scopoletin (ng/mg fresh sample) in *Arabidopsis* roots and leaves +/- standard deviation as determined by LC-MS from at least four samples (n≥4).

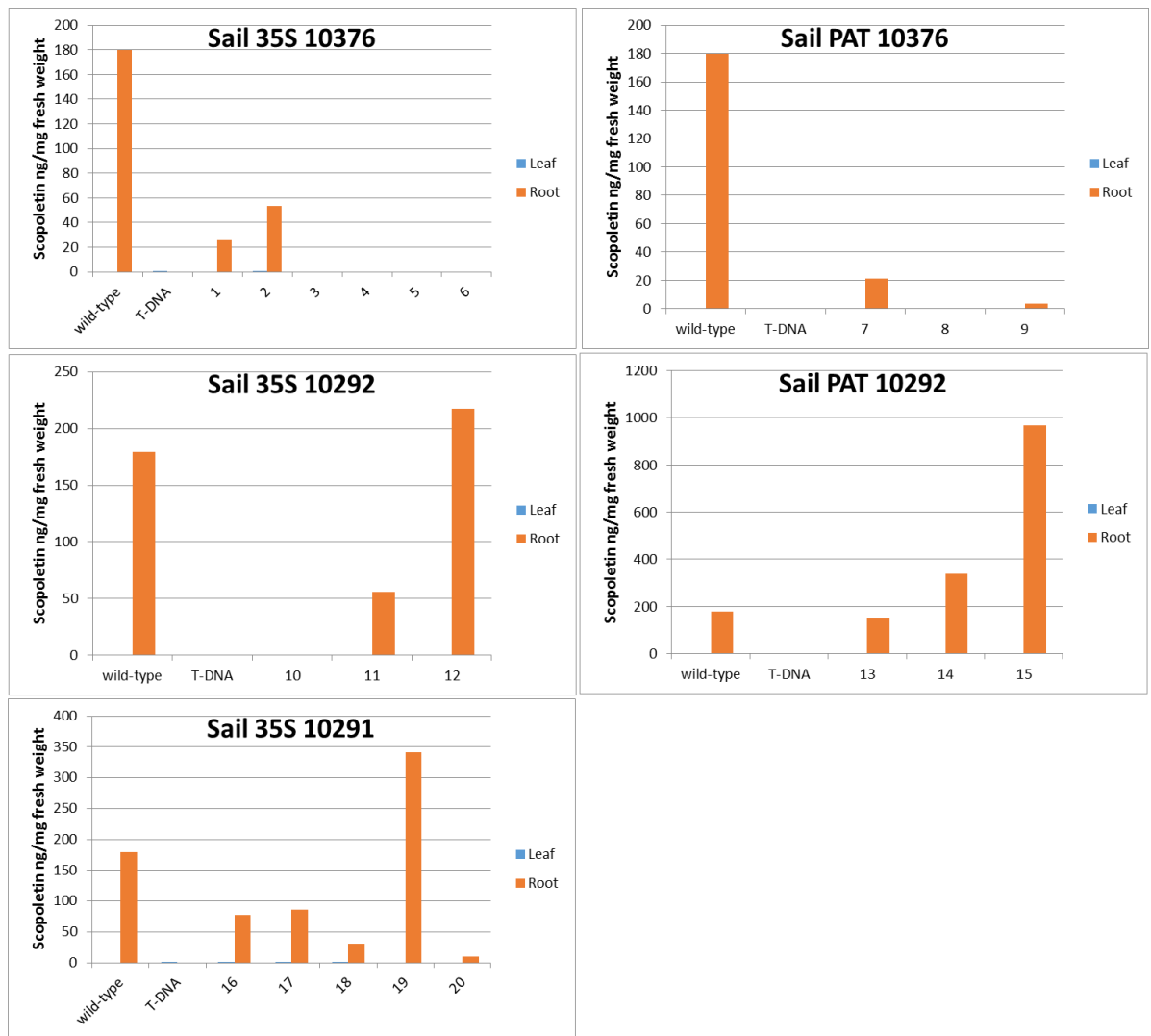
Compared to data from the literature, our scopoletin levels were surprisingly high. The wild-type *Arabidopsis* root scopoletin level was approximately 2 ng/mg fresh weight from Kai et al. and approximately 3.4 ng/mg fresh weight from Siwinska et al.; which are much lower than that of our results (Kai et al., 2008, Siwinska et al., 2014). In addition, the data from each individual sample showed large variation, which gave the mean scopoletin level a large standard deviation (error bar in figure 4.7). Despite the difference, the pattern of the data was the same. Scopoletin accumulated in wild-type *Arabidopsis* roots but not in the leaves, and in T-DNA insert mutants scopoletin was accumulated in neither the roots nor the leaves (Kai et al., 2008, Kai et al., 2006).

The difference in scopoletin level may be due to the plant growth and sample processing. Due to a limit of space and instruments, hydroponics (as used by

Siwinska et al.) was not possible, so the *Arabidopsis* plants were grown on agar plates, which caused difficulties in harvesting fresh and clean root samples. Although not stated in the papers mentioned above, different methods of sample processing may also contribute to this difference (Kai et al., 2006). In *Arabidopsis* roots, the levels of scopolin (the glycoside of scopoletin) were approximately 85-fold higher than scopoletin (Kai et al., 2006). Since the root samples were stored and processed over a long period of time, the spontaneous slow decomposition of scopolin to scopoletin would lead to an increased level of the latter in the samples (Al-halaseh, pers. Comm.). The *Arabidopsis* root samples for quantification were harvested, stored and processed over a time-span of two years. The different instruments and their set-up, as well as the sample preparation may also contribute to the variation of the data.

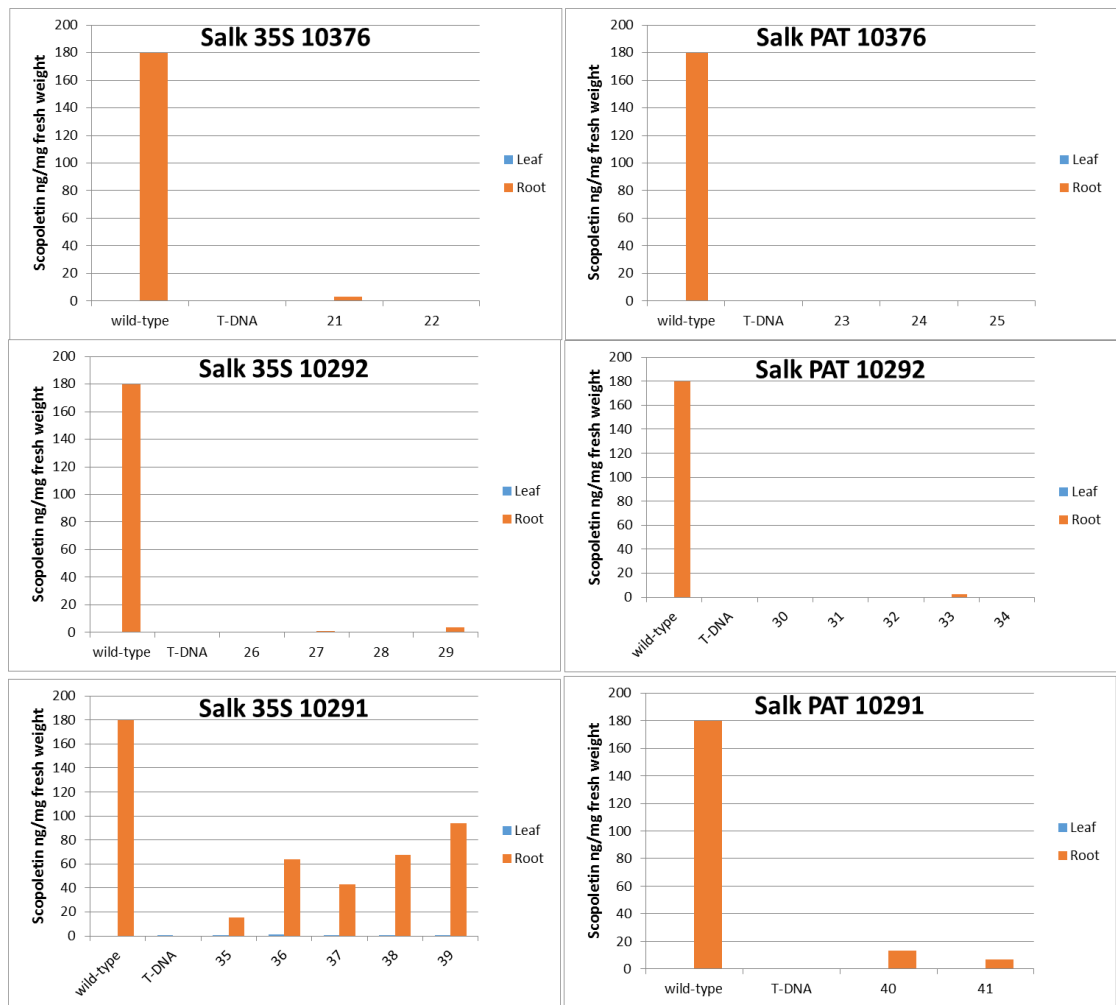
#### **4.2.5 Confirmation of the functional identity of cassava candidate genes**

Scopoletin levels in the roots of homozygous *Arabidopsis* F6'H1 T-DNA insert mutants that were transformed with the three cassava candidate genes that were regulated by either a root-specific *StPAT* promoter or a constitutive CaMV 35S promoter were compared to untransformed control wild-type and mutants plants as shown in figure 4.8 and 4.9.



**Fig 4.8 The scopoletin biosynthetic complementation in Sail\_1252\_A10 F6'H1 T-DNA insert mutants of *Arabidopsis* with different cassava candidates and promoters.** Data are mean levels of scopoletin (ng/mg fresh sample) in *Arabidopsis* roots and leaves as determined by LC-MS. Every histogram shows scopoletin levels of different lines in one combination of *Arabidopsis* F6'H1 T-DNA insert line, promoter, and cassava gene, compared to wild-type and F6'H1 T-DNA insert lines. Sail: Sail\_1252\_A10 T-DNA insert line; 35S: CaMV 35S promoter; PAT: StPAT promoter; 10291, 10292, and 10376: the cassava gene over-expressed. Wild-type: wild-type *Arabidopsis* samples. T-DNA: Samples of Sail\_1252\_A10 T-DNA insert line without cassava F6'H genes. Numbers from 1 through 20 indicate different transformant lines. Notice that in most samples leaf scopoletin levels are too low to be detected. Data of the wild-type and T-DNA *Arabidopsis* come from at least 4 biological replicates, and data of each transgenic line (labelled with a number from 1 to 20) are from two biological replicates.





**Fig 4.9 The scopoletin biosynthetic complementation in Salk\_129938 F6'H1 T-DNA insert mutants of *Arabidopsis* with different cassava candidates and promoters.** Data are mean levels of scopoletin (ng/mg fresh sample) in *Arabidopsis* roots and leaves as determined by LC-MS. Every histogram shows scopoletin levels of different lines in one combination of *Arabidopsis* F6'H1 T-DNA insert line, promoter, and cassava gene, compared to wild-type and F6'H1 T-DNA insert lines. Salk: Salk\_129938 T-DNA insert line; 35S: CaMV 35S promoter; PAT: StPAT promoter; 10291, 10292, and 10376: the cassava gene over-expressed. Wild-type: wild-type *Arabidopsis* samples. T-DNA: Samples of Salk\_129938 T-DNA insert line without cassava F6'H genes. Numbers from 21 through 41 indicate different transformant lines. Notice that in most samples leaf scopoletin levels are too low to be detected. Data of the wild-type and T-DNA *Arabidopsis* come from at least 4 biological replicates, and data of each transgenic line (labelled with a number from 21 to 41) are from two biological replicates.

Scopoletin accumulation in *Arabidopsis* roots clearly demonstrated that transformants of all three cassava candidate genes regulated by either of the two promoters were capable of complementing the scopoletin synthesis in homozygous F6'H1 T-DNA insertion mutations in both the Salk\_129938 and Sail\_1252\_A10 *Arabidopsis* lines. However, it was interesting that different transformant lines with the same promoter-gene combination led to significantly different scopoletin levels in their roots, including some in which scopoletin was not detectable. An interesting fact was that transgenic lines based on T-DNA insert mutant Salk\_129938 showed lower scopoletin accumulation regardless of the gene and promoter used when compared to lines based on mutant Sail\_1252\_A10. It was unclear why such a difference existed, but the different T-DNA insert mutants should not affect the ability of the cassava candidate genes. Since in Salk\_129938 mutant the F6'H1 gene was still transcribed (although with the T-DNA insert that disabled its translation), the mRNA of F6'H1 gene may have interacted with the external cassava F6'H1 candidate genes, inhibiting their expression due to the similarity in sequences (Hobbs et al., 1993, Jorgensen, 1990). Gene 10376 showed lower ability of scopoletin synthesis compared to gene 10291 and 10292, although it did show some F6'H1 activity with both the CaMV 35S and StPAT promoters. Gene 10291 led to a relatively stable scopoletin synthesis in the Sail\_1252\_A10 and surprisingly also in the Salk\_129938 mutant when being regulated by a constitutive CaMV 35S promoter, but with a root-specific StPAT promoter it did not show as high scopoletin synthesis. Gene 10292 showed similar scopoletin synthetic levels in both constitutive CaMV 35S promoter and root-specific StPAT promoter driven plants. Although it led to very high scopoletin levels in a few individuals, the scopoletin accumulation in different individuals varied a lot.

However, this variation in expressing levels between different transgenic lines is not unexpected, as, due to the random nature of T-DNA insertion, every single transformant has a unique and unpredictable 'context' of chromatin structure and flanking sequences (De Bolle et al., 2003). This can give each transgenic line a unique expression strength. The constitutive promoter CaMV 35S has been shown to exhibit highly varied expression level in different transformants (De Bolle et al., 2003). Although there has been no extensive study on the StPAT promoter, there are data showing the variable expression level of transgenes under the control of

this promoter (Rocha-Sosa et al., 1989). Additionally, although the expression of single insertions can be unpredictable, lines with multiple inserts in the host genome can be even more so. For example, in tobacco multiple copies of the transgene were shown to either increase or decrease the expression of a target gene (Hobbs et al., 1993), or even silence its expression (Jorgensen, 1990) due to RNA interference-related mechanisms (Napoli et al., 1990). To avoid such homology silencing, *Arabidopsis* transformant lines with only single copy of insert were selected as discussed above. An additional contribution to the high variation measured within the samples could be due the challenges of sample preparation from small numbers of roots and the limited numbers of replicates tested.

Due to the variations outlined above, it did not prove possible to determine statistically whether a gene driven by a constitutive (CaMV 35S) or a root-specific (*StPAT*) promoter led to a more efficient expression in the root. In addition, it was not clear which of these three cassava candidate genes showed the highest feruloyl CoA 6'-hydroxylase activity in transformed *Arabidopsis*. However, the impression from the result was that gene 10292 led to the highest scopoletin levels while gene 10376 led to the lowest ones.

### 4.3 DISCUSSION

This chapter discusses how the expression of the cassava F6'H1 candidate genes in *Arabidopsis* F6'H1 T-DNA insert mutants enables the verification of the function of the cassava candidates. The three tested genes enabled *Arabidopsis* F6'H1 T-DNA insert mutants to accumulate scopoletin proving that at least these three members of the cassava F6'H1 gene family are capable of catalysing feruloyl CoA 6'-hydroxylation in the scopoletin biosynthetic pathway.

The results of the complementation of *Arabidopsis* scopoletin biosynthesis should have helped us choose the gene for cassava RNAi construct. Due to the slow growing nature of cassava, a decision was made at the beginning of the project before the results of the expression of cassava candidates in *Arabidopsis* F6'H1 T-DNA insert mutants were available. Gene 10376 was selected as the target of RNAi constructs due to its closest similarity to the coding sequence of *Arabidopsis* F6'H1 (figure 3.3) and a strong evidence of expression in cassava roots (table 3.5). We only chose a single gene (10376) for cassava RNAi because of limited time and

cassava transformation facilities. However, with the data of scopoletin complementation in *Arabidopsis* F6'H1 T-DNA insert mutants, gene 10376 showed limited ability of complementing the of F6'H1 T-DNA insert mutants in *Arabidopsis*, though whether this was due to the gene *per se*, or to the position of insertion in the *Arabidopsis* genome of the few transgenic lines tested, is uncertain. With enough time (more than a three year PhD project) and opportunities, we would have expressed every cassava F6'H1 candidate in *Arabidopsis* F6'H1 T-DNA insert mutants and analysed a large number of independent single-insert lines from each gene-construct in order to identify the best candidate for the basis of an RNAi construct for cassava. Nonetheless, an RNAi construct based on gene 10376 is likely to target the entire cassava F6'H1 gene family, as figure 3.3 shows the similarity between these cassava candidates. In addition, there is a highly conserved sequence of 22 nucleotides shared between all these cassava genes, which should act as a good target of RNA silencing. The mechanism of RNAi by external double stranded RNA (dsRNA) is still unclear but there is evidence that it is related to the post-transcriptional gene-silencing (PTGS) involved in viral defencing and epigenetic gene expression regulation mechanism (Elbashir et al., 2001, Klahre et al., 2002). In plants, the PTGS mechanism requires small antisense RNA (approximately 25-nt) complementary to its target mRNA sequence (Hamilton, 1999). In *Arabidopsis* it is likely that the external dsRNA from the RNAi construct is processed to 21-25 nucleotides small interfering RNA (siRNA) and assembled into the RNA-induced silencing complex (RISC) and degrades the target mRNAs having complementary sequences (Gazzani et al., 2004, Qi et al., 2005). The conserved sequences in the cassava F6'H1 candidate gene family should be targeted by the siRNAs from the sequence of gene 10376. As a summary, although gene 10376 itself may not be the gene performing highest F6'H1 activity, an RNAi structure based on its sequence should also target the other genes. The selection of the target gene for RNAi to down-regulate in cassava will lead to the cassava transformation and PPD quantification, which will be discussed in chapter 5.

## **5 DOES INHIBITING 6'-HYDROXYLATION OF FERULOYL COA IN CASSAVA ROOTS ALTER THE BIOSYNTHESIS OF SCOPOLETIN AND PPD RESPONSE?**

### **5.1 INTRODUCTION**

#### **5.1.1 Post-harvest physiological deterioration (PPD)**

In the previous chapter, one cassava F6'H1 candidate gene (10376) was chosen as the basis on which to construct RNAi vectors for transformation into cassava. In this chapter we describe this process and present the data derived from expressing such vectors on the PPD response of cassava roots.

PPD is a phenomenon unique to cassava, which causes the cassava roots to deteriorate and lose palatability in a few days after harvest (Blagbrough et al., 2010). This process is triggered by the wounding inevitably caused by harvesting and handling, and it occurs within 24 hours during which a cascade of physiological biochemical responses are activated (Reilly et al., 2007). During the PPD, a blue-black discoloration develops in the roots due to the accumulation of phenolic compounds and their oxidation by reactive oxygen species (ROS). Amongst these phenolic compounds, scopoletin, a hydroxycoumarin, shows the most significant accumulation (from less than 1 ng/mg fresh weight to approximately 80 ng/mg fresh weight in 2-3 days) (Wheatley and Schwabe, 1985). Although this dramatic accumulation of scopoletin is not the only event occurring during PPD, its oxidation by peroxidase and H<sub>2</sub>O<sub>2</sub> to form a blue-black pigment certainly plays a major role in the discoloration of the vascular tissue of cassava root (Beeching et al., 2002). To study the role of scopoletin in PPD development, we aim to block the accumulation of scopoletin by inhibiting an important step (cassava F6'H1 candidate-mediated feruloyl CoA 6'-hydroxylation) in its biosynthesis (Bayoumi et al., 2008a, Kai et al., 2008).

Feruloyl CoA 6'-hydroxylase 1 (F6'H1) is crucial to the biosynthesis of scopoletin in *Arabidopsis* roots, and in its absence scopoletin cannot be produced and does not accumulate in the roots (Kai et al., 2008, Kai et al., 2006). F6'H1 converts

feruloyl-CoA to 6'-hydroxyferuloyl-CoA in the dominant biosynthetic pathway of scopoletin in *Arabidopsis* roots (pathway 3, figure 1.11). In the genome of *Arabidopsis*, there are two feruloyl CoA 6'-hydroxylases, namely F6'H1 (Phytozome ID: At3g13610) and F6'H2 (Phytozome ID: AT1G55290). However, the expression of F6'H2 is negligible in mature *Arabidopsis* roots compared with that of F6'H1 (approximately one fifteenth) (Winter et al., 2007), and if F6'H1 is inactivated by a T-DNA insertion mutation, scopoletin accumulation in *Arabidopsis* roots is below detectable levels. Therefore, the 6'-hydroxylation of feruloyl CoA, and thus scopoletin biosynthesis in *Arabidopsis* roots, hinges on only one gene, F6'H1. In cassava, seven F6'H1 candidate genes were identified as discussed in chapter 3. It is unclear which or how many genes in this gene family of F6'H1 candidates regulate 6'-hydroxylation of feruloyl CoA in cassava, as the F6'H1 candidates in the gene family are highly similar to each other, as discussed in chapter 3. However, at least three genes (10291, 10292, and 10376) were confirmed as being able to functionally complement *Arabidopsis* F6'H1 T-DNA insertion mutants. Therefore, in cassava roots feruloyl CoA 6'-hydroxylation is likely to be regulated by multiple genes. Gene 10376 was chosen as the basis on which to construct RNAi vectors as sequences within it were sufficiently highly similar to those within other family members for such constructs to have the potential to down-regulate the expression of several genes in the gene family simultaneously. Such a general inhibition of cassava F6'H1 candidates may prevent the 6'-hydroxylation of feruloyl CoA, thereby reducing the synthesis of scopoletin in cassava roots, so as to provide an insight into the role of scopoletin in PPD.

### **5.1.2 RNAi inactivated cassava gene expression**

The interference phenomenon of external double-strand RNA (dsRNA) was first observed in *Caenorhabditis elegans*, and later recognised in plants as a defensive mechanism against viruses (Fire et al., 1998, Waterhouse et al., 2001). This endogenous gene-silencing mechanism is referred to as post-transcriptional gene silencing (PTGS) and in plant organisms the dsRNA is processed to approximately 25 nt small RNA fragments to target and degrade its complementary mRNA target (Hamilton, 1999).

In plants, a well-established method to silence an endogenous gene is to introduce a fragment from the gene with its inverted repeat that can ‘fold back’ to form a double-stranded RNA molecule (figure 5.1 D). The silencing efficiency is enhanced if the repeats are separated by an intron, which can be prepared conveniently in a pHannibal or pKannibal vector with the intron built-in (figure 5.1 B) (Helliwell and Waterhouse, 2003). This construct produces a ‘hairpin’ dsRNA once transcribed, which leads to approximately 60% of the individuals expressing it to have the target-gene down-regulated (Smith et al., 2000). In this method a 300 to 500 bp fragment from the target gene is required, and the position of this fragment in the gene is not critical to a successful gene silencing (Wesley et al., 2001). The RNAi method is known to be able to knock out either multiple copies of the same gene, or multiple genes in a gene family, with only one RNAi transgenic insertion (Preuss and Pikaard, 2004), which makes it easier for us to study the collective function of the cassava F6'H1 candidate gene family.

## **5.2 AIMS**

The aims of this chapter are:

- 1, Prepare the selected cassava F6'H1 candidate gene (gene 10376) for RNAi;
- 2, Transform wild-type (TMS 60444) cassava to express the RNAi construct;
- 3, Grow and harvest the transgenic RNAi cassava to obtain transgenic roots;
- 4, Study and quantify PPD responses of wild-type and RNAi transgenic cassava roots.

## **5.3 RESULTS**

### **5.3.1 Cassava candidate gene 10376 was used to construct RNAi**

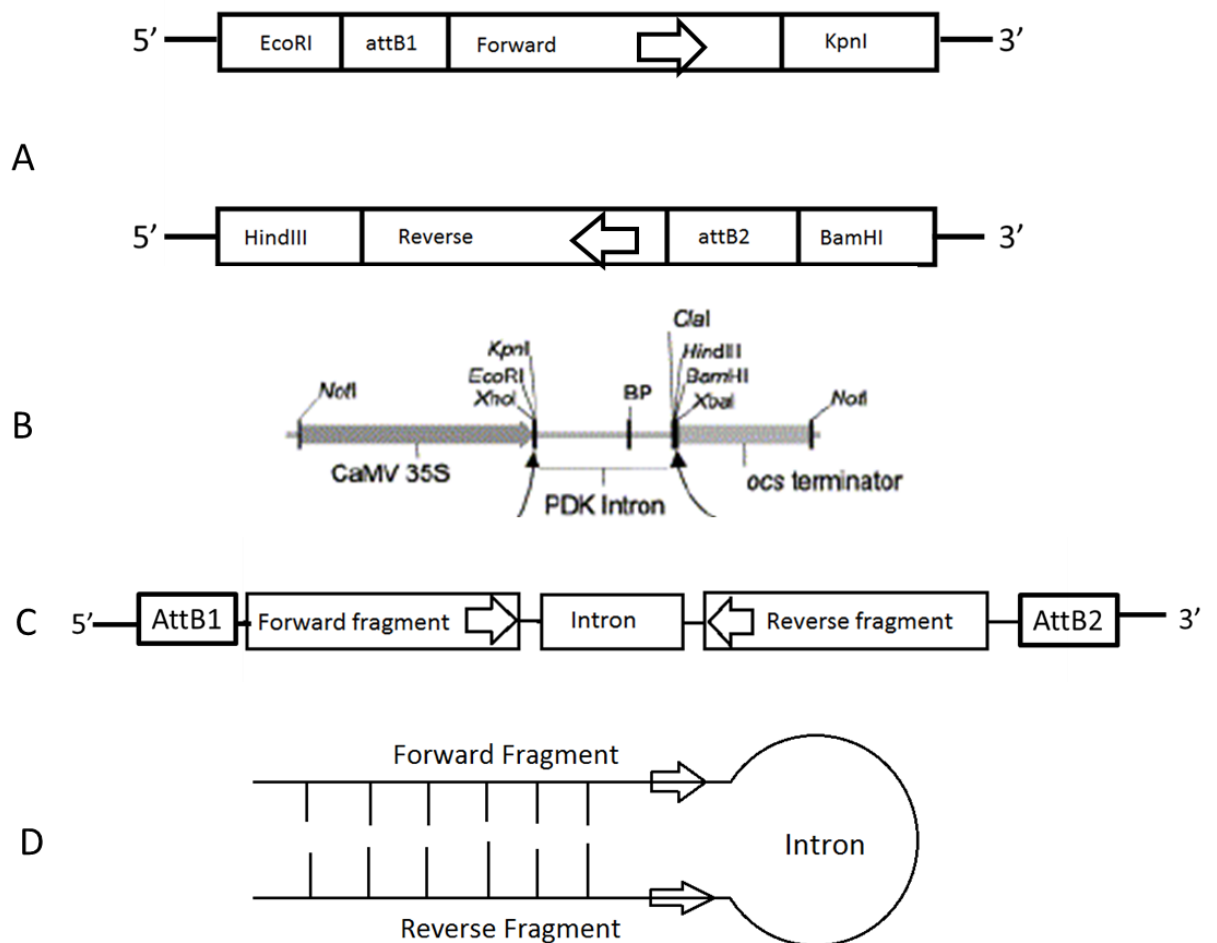
The first exon of cassava gene 10376 was amplified with primer pair 18 to obtain a 471 bp fragment of the gene. This fragment was either further amplified with primer pairs 19, 20, and 21 to produce the ‘forward’ fragment, or pairs 22, 23, and 24 for the ‘reverse’ fragment, consecutively to add appropriate AttB sites and restriction sites to it. The structures of the two fragments are shown in figure 5.1 (A). In the

forward and reverse fragments, the 471 bp sequences from gene 10376 were in reverse-complement orientation. The forward and reverse fragments were digested with appropriate restriction enzymes (EcoRI and KpnI for the forward fragment, and HindIII and BamHI for the reverse fragment) and inserted into the pKannibal vector that was digested with the same restriction enzymes (Helliwell and Waterhouse, 2003) so that a structure as in Fig. 5.1 (C) was constructed. The forward and reverse fragments flanking a PDK intron (named the RNAi cassette) were cloned into the Gateway<sup>®</sup>-competent pCAMBIA 1305.1 vectors as described in chapter 4 via Gateway<sup>®</sup> cloning. The expression of the RNAi cassette was regulated by either a constitutive CaMV 35S promoter or a root-specific *St*PAT promoter in the Gateway<sup>®</sup>-competent pCAMBIA 1305.1 vector. The RNAi vectors were sequenced after every step so as to ensure that the final constructs had the correct sequence.

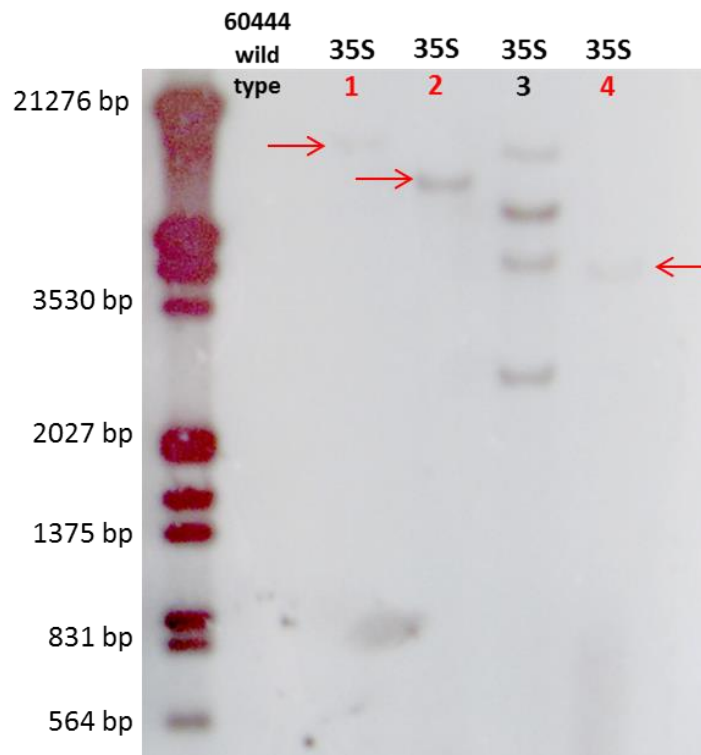
### 5.3.2 Wild-type cassava was transformed

Due to time and logistical constraints the transformation of wild-type cassava (TMS 60444) and the verification of transgenic lines were generously carried out at ETH-Zurich (where the experience and facilities were available) by Dr Ima M. Zainuddin. Transgenic cassava lines carrying the RNAi construct (as a T-DNA insert mutant) regulated by either constitutive CaMV 35S promoter or *St*PAT promoter were generated via *Agrobacterium*-mediated FEC transformation as described in chapter 2. To select transformant lines carrying only one copy of the RNAi T-DNA insert mutant, the genomic DNA of the RNAi transgenic cassava lines were tested with Southern blotting. Figure 5.2 shows a picture of Southern blotting film, on which transformant lines with only one copy of insert were detected. From a large number of transformant lines, nine lines with only one copy of T-DNA insert were identified. Six lines of the single-insert transformants were regulated by the constitutive CaMV 35S promoter and three by the root-specific *St*PAT promoter. These lines are referred to as 35S-1, 35S-2, 35S-4, 35S-5, 35S-6, 35S-L, PAT-8, PAT-11, and PAT-14. The 35S or PAT in a name indicates the promoter that regulates the transcription of the RNAi constructs (35S for CaMV 35S, and PAT for *St*PAT).





**Fig. 5.1 The RNAi vector produces a hairpin structure generated by forward and reverse fragments of the target sequence.** (A) the appropriate restriction sites (EcoRI and KpnI in forward fragment, and HindIII and BamHI in reverse fragment) and the attB sites are engineered into the target sequences; (B) two polylinkers and a PDK intron in the pKannibal vector facilitate engineering of the RNAi vector. Picture modified from (Helliwell and Waterhouse, 2003); (C) the forward-and-reverse fragments with an intron in between is engineered in the pKannibal vector; (D) the transcribed single-strand RNAi sequence spontaneously forms a hairpin structure.

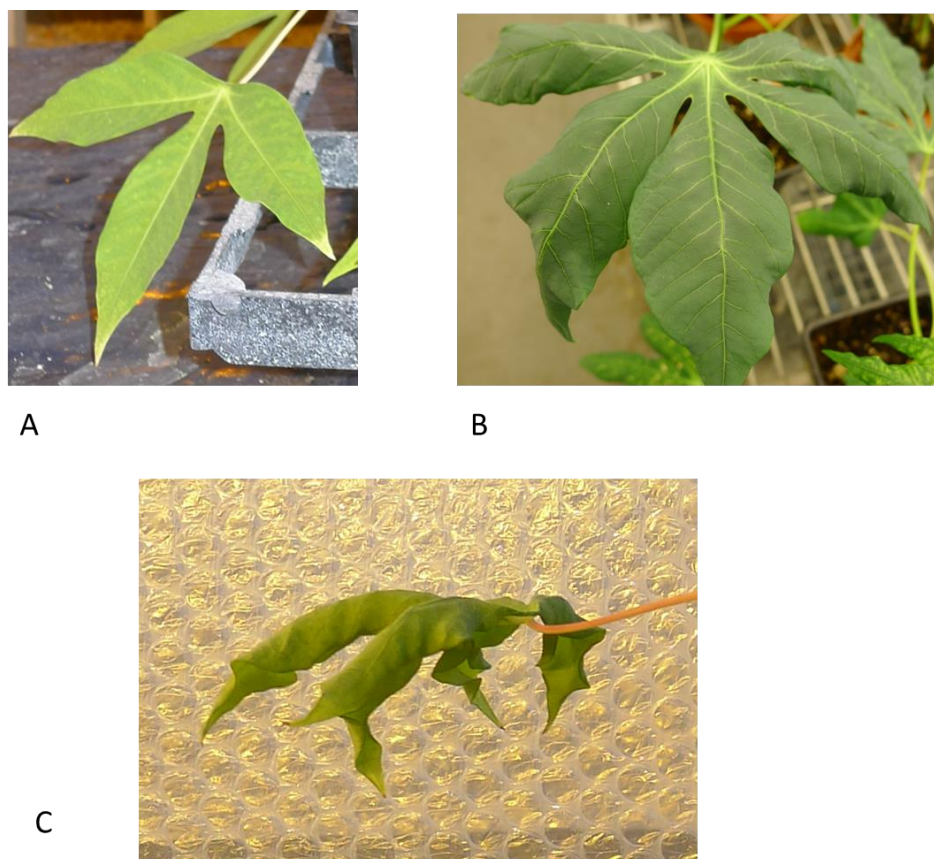


**Fig. 5.2 Number of bands in a Southern blotting test indicates the copy numbers of T-DNA inserts in the transgenic lines.** From left to right: the DNA markers and their size; the wild-type plant (TMS 60444) showing no band; line 35S-1, 35S-2 and 35S-4 show a single band (indicated with red arrowheads), and line 35S-3 shows multiple bands. Picture modified from Zainuddin.

It is important to obtain several distinct single-insert transformant lines from the transformation as the strength of the expression of a transgene in a transformant line is influenced by the location where it is inserted into the chromosome and its flanking sequences (i.e. the epigenetic 'context') (Matzke and Matzke, 1998). Due to the random nature of *Agrobacterium*-mediated T-DNA insertion, the location of the integrated T-DNA transgene (in this case the RNAi construct) in the host's genome is unpredictable (Alonso et al., 2003, Hiei et al., 1994). Therefore, the location and context of the T-DNA insert in each distinct transformant line is different, as may be its expression. As a result, the same transgene may have dramatically distinct expression behaviours in different transformant lines, or even be silenced (Peach and Velten, 1991). The different single-insert transformant lines must be screened to identify the strengths of their gene expression, so as to obtain transformant lines with robust expression of their transgenes (i.e. the RNAi construct).

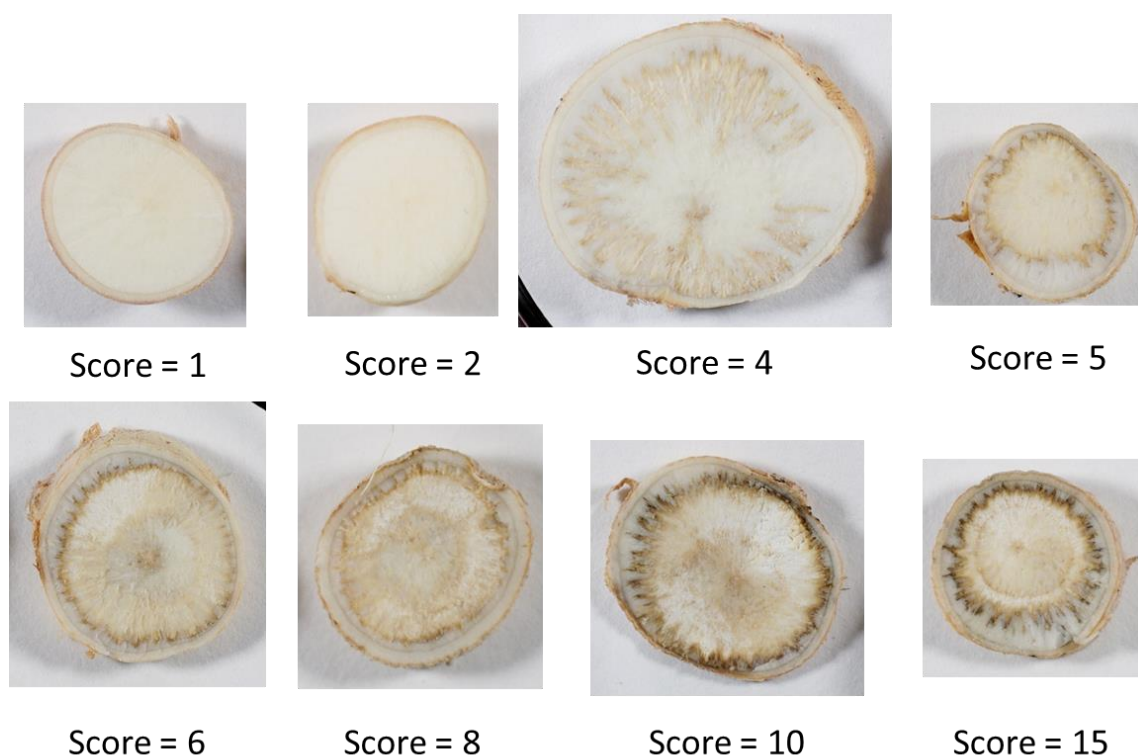
### 5.3.3 RNAi altered cassava phenotypes and PPD discoloration

The nine single-insert RNAi transformant lines were multiplied *in vitro* to form plantlets and six to ten individual plantlets of each transformant line, as well as plantlets of wild-type cassava (TMS 60444), were allowed to grow for six to seven months in the University glasshouse. The RNAi transgenic individuals exhibited altered phenotypes compared to the wild-type. The leaf lobes of RNAi plants often grew broader and their tips tended to have a ‘folded’ or even ‘helix’ shape pointing downwards (figure 5.3). It is unclear why the transgenic lines have such an altered phenotype. A probable reason may be that either the off-target effect of the RNAi constructs or the *Agrobacterium*-mediated transformation itself, including the passage through tissue culture, affects expression of genes related to leaf formation. A way to shed light on this, though not performed, would be to grow subsequent generations of plants from conventional stem cuttings.



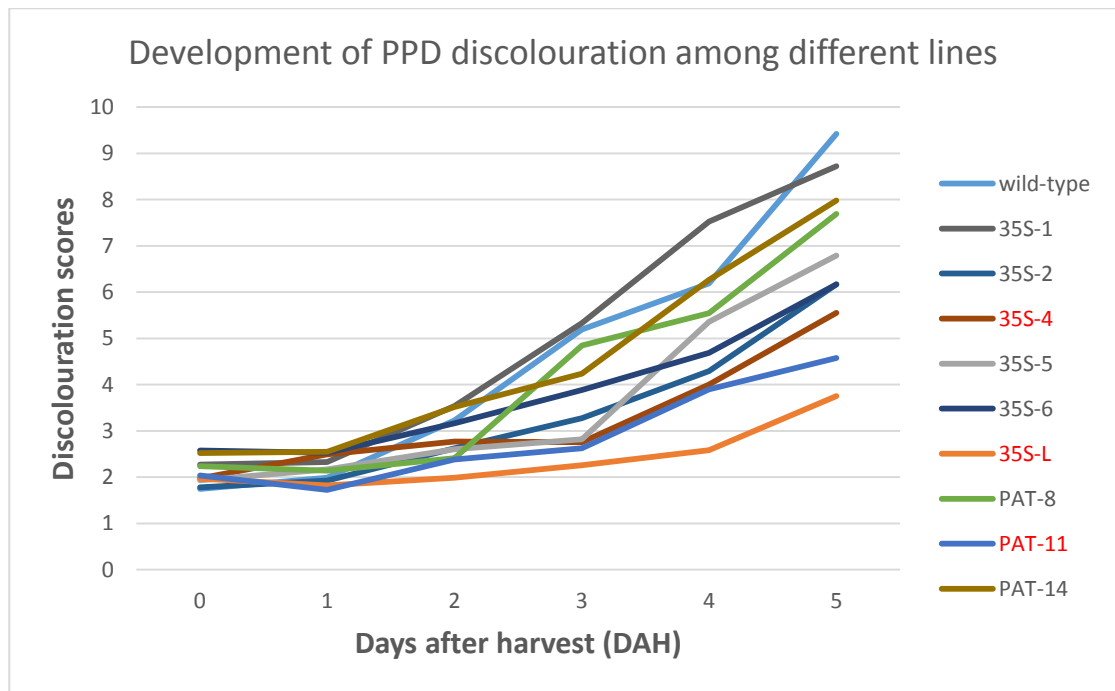
**Fig. 5.3 Wild-type and RNAi transgenic cassava have different phenotypes.** (A) wild-type cassava leaf lobes were flat and narrow; (B) RNAi transgenic cassava leaf lobes were broader; (C) RNAi transgenic cassava leaf lobes were folded and pointing downwards.

The roots of both wild-type and RNAi transgenic cassava plants were harvested after 6-7 months growth and their PPD discoloration was induced and evaluated as described in chapter 2. The freshly harvested roots (0 DAH) through a time-course of 5 days were photographed and the discolorations were quantified as discussed in Section 2.10.2. Basically the discoloration score of a sample indicates how many folds its greyscale is darker than that of the filter paper (Whatman). Note, that because the filter paper is used to standardise the measurements a root at time zero, despite not suffering any PPD, will have a score of above zero. A score below 2 implies no discoloration, and a score above 6 implies significant discoloration. Figure 5.4 shows examples of roots of different scores of discoloration. The results are shown in figure 5.5.



**Fig. 5.4 Examples of cassava root slices at different stages of discoloration.**

The score was calculated by dividing the greyscale reading of the parenchyma part of the roots by that of the filter paper beneath. The shaded area of filter paper was avoided.



**Fig. 5.5 The discoloration development of the wild-type cassava compared with that of the nine RNAi T-DNA insert mutants during PPD.** This chart shows the relative discoloration of samples in a time-course of 5 days as described in 5.3.3. The names of the three transformant lines showing the most significant discoloration alleviation are shown in red in the figure legends. Data are mean levels of discoloration scores from at least 6 samples and error bars are omitted for clarity. (n≥6)

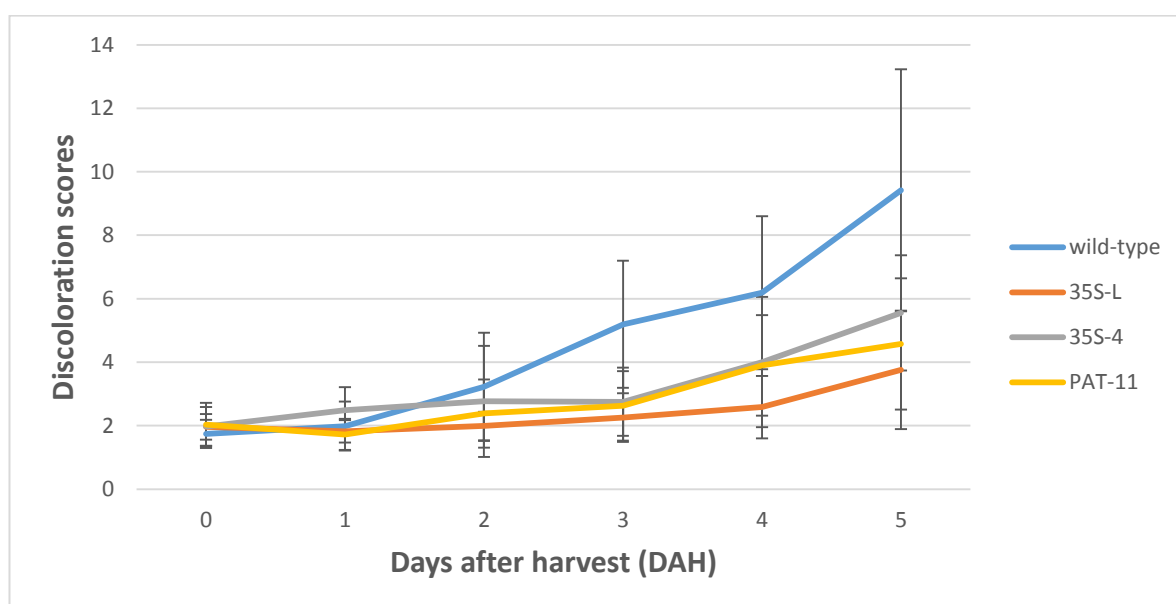
The RNAi transgenic cassava lines showed wide range of discoloration levels, with both the constitutive CaMV 35S and the root-specific *StPAT* promoters leading to effective discoloration alleviation in certain lines. Such a contrasting response of independent single insert lines is common in transformation experiments, for which it is important to screen several different lines. The insertion of a transgene into the genome is essentially random and can equally into a region that is transcriptionally inactive as one that is transcriptionally active. Therefore, the expression levels of the transgene and the resultant phenotype, in this case the discoloration level, can vary considerably between independent lines (Matzke et al., 2001, Peach and Velten, 1991). For example, in line PAT-11, on 5 DAH the lowest discoloration reading was 1.71 (similar to that of wild-type fresh roots) and the highest reading was 7.52 (the average level of wild-type roots on 4-5 DAH). For these reasons as many transgenic lines as is convenient are usually screened and the best

performing ones are selected to forward for further analysis. Therefore, the three RNAi transformant lines showing the most significant discoloration alleviation (namely, 35S-4, 35S-L, and Pat-11) were chosen to be taken forward.

Of the three best performing lines, 35S-4, 35S-L, and PAT-11, two were driven by the constitutive CaMV 35S promoter, and one by the root-specific *St*PAT promoter. Therefore, these lines should enable the comparison of cassava plants in which cassava F6'H1 candidates were inhibited in the roots alone (*St*PAT) and in the whole plant (CaMV 35S). Removing the low performing lines enables a clearer view of the discoloration between the selected transgenics and the wild-type roots (figure 5.5 and 5.6). As discussed previously, the data derived from all lines, either transgenic or wild-type, are noisy and fluctuating. The standard deviations of these lines are great (error bars, figure 5.6), which indicate that samples harvested from the same line at the same stage vary dramatically. It is important to remember that the cassava plants were grown in 9 cm x 9 cm pots for six to seven months before harvest, which means that although the behaviour of these roots is likely to be similar to that of field-grown mature storage roots, it may not be identical. In other words, the sample roots may be immature and their PPD behaviour may not be as reliable as that of the mature storage roots. In addition, although the total time that every plant spent in the pot is similar (6-8 months), it is impossible to know the exact time a storage root had grown since the beginning of its enlargement. As evidence, the size and maturity stage of different root samples also vary considerably. For example, the diameter of the largest wild-type sample slice was approximately 3 cm but the smallest sample was around 1 cm. Although the extremely immature samples had been omitted from the statistics, different stages of maturity may still contribute to the unstable discoloration behaviour within a certain transgenic line.

However, regardless of the variation of the data, the discoloration of the three selected RNAi transformant lines showed significant differences when compared to that of the wild-type line in a t-test (table 5.1). The discoloration data of each individual transgenic line at 3 DAH, 4 DAH, and 5 DAH were compared with that of the wild-type. All the three transgenic lines showed P values less than or around 0.01 in all stages, which indicated that there were 99% confident that all the three transformant lines showed significantly alleviated discoloration in the advanced stages of PPD. These results confirmed that the RNAi transformation delayed PPD

discoloration successfully, and these three RNAi transgenic lines were also used in downstream analyses.



**Fig. 5.6 The discoloration development of the three selected RNAi T-DNA insert mutants compared with the wild-type cassava.** This chart shows the relative discoloration of samples in a time-course of 5 days as described in 5.3.3. Data are mean discoloration scores +/- standard deviation from at least 6 replicates. (n≥6)

P values	Wild-type 3DAH	Wild-type 4DAH	Wild-type 5DAH
<b>35S-4</b>	0.0010	0.0148	0.0117
<b>35S-L</b>	0.0002	0.0001	0.0002
<b>PAT-11</b>	0.0001	0.0021	0.0002

**Table 5.1 T-tests reveal the significant discoloration alleviation in the three transformant lines.** Numbers in the table indicate the P values of the discoloration data compared with wild-type at 3, 4, or 5 days after harvest. P-values between 0.05 and 0.01 (95% confidence) are shown in blue and values less than 0.01 (99% confidence) are shown in red.

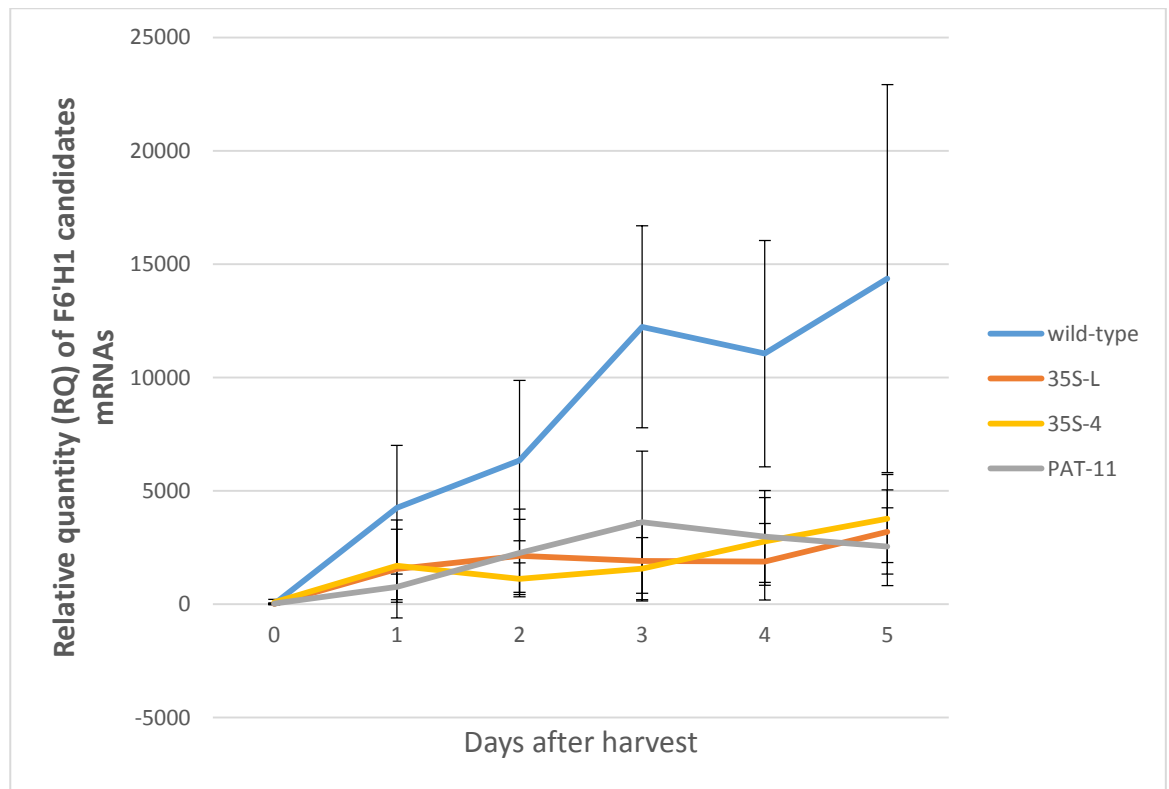


#### 5.3.4 RNAi down-regulated the expression of the cassava F6'H1 candidates

As discussed above, root samples from wild-type cassava and the three chosen RNAi transformant lines were allowed to develop PPD over a time-course of 5 days after harvest. A root slice was photographed for PPD quantification and the same slice was also processed for RNA extraction and scopoletin quantification. RNA of these root slices were extracted and used in quantitative RT-PCR as described in chapter 2. A cassava ubiquitin gene (Phytozome ID: cassava4.1\_009650m.g, locus name: Manes.07G019300) was used as the reference gene (Moreno et al., 2011). Due to the extreme similarity between the cassava F6'H1 candidates, it proved impossible to design primers for quantitative RT-PCR that could with certainty differentiate between the individual candidates. Therefore, a pair of primers (primer pair 12) capable of amplifying multiple genes was used. As a result, in all probability, the primers amplified cDNAs derived from the mRNA of all four related genes, 10291, 10292, 10376, and 27567, thereby providing an overall measure of feruloyl CoA 6'-hydroxylase gene activity in post-harvest cassava storage roots (figure 5.7).

At harvest F6'H1 candidates mRNAs were virtually below detectable in all roots (approximately 1/500 of the signal strength of the housekeeping reference gene ubiquitin in qRT-PCR), for both wild-type and transgenic. Over the 5-day time-course monitored the F6'H1 candidates' mRNA increased in all plants. However, there is a significant difference between levels reached in the wild-type and the three RNAi transgenics. The former achieving levels five-fold greater than the latter, thereby confirming that the RNAi construct, while not completely preventing, is capable of significantly down-regulating the expression of the cassava F6'H1 candidate gene family, as shown in table 5.2. This result is not surprising since dsRNA-related gene-silencing methods often lead to partial gene inactivation, and multiple genes in the targeted F6'H1 gene family may also contribute to incomplete gene knock-out (Coleman et al., 2008, Helliwell and Waterhouse, 2003, Wesley et al., 2001). The significant reduction, but not complete inhibition, of the total expression of genes in the cassava F6'H1 candidate gene family by the RNAi construct may lead to a lowered but not fully inhibited scopoletin accumulation in the cassava roots.





**Fig. 5.7 The expression levels of the F6'H1 candidate gene family in wild-type cassava and three selected RNAi transformant lines.** Data are mean relative quantification of cassava F6'H genes mRNA  $\pm$  standard deviation in wild-type and transgenic cassava over a time-course of five days as determined by qRT-PCR from at least four samples ( $n \geq 4$ ).

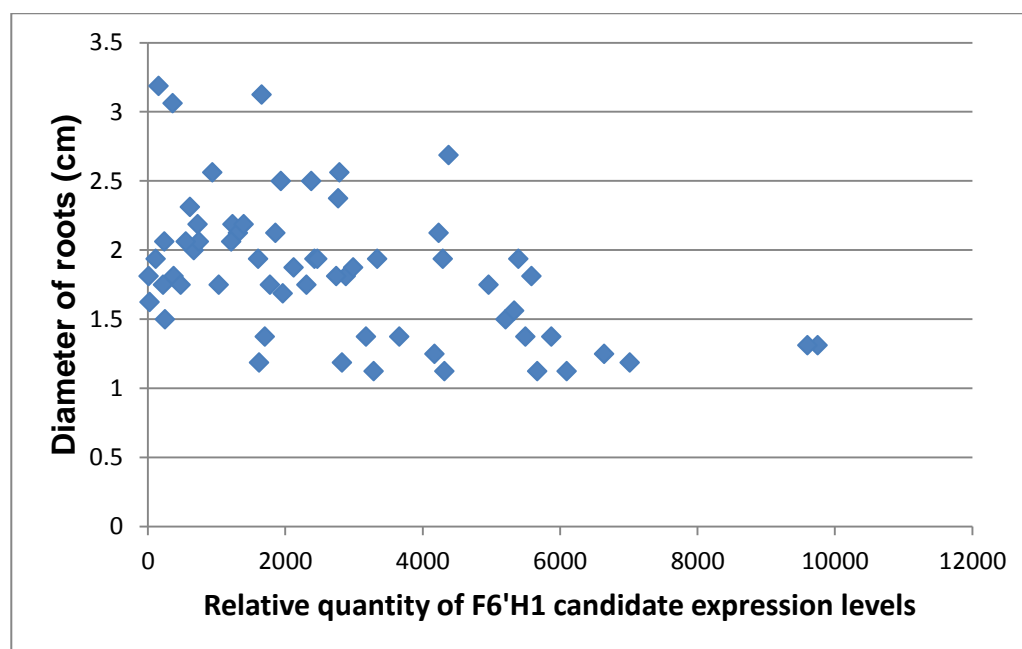
Another observation was that sizes of the roots also contributed to the diverse levels of F6'H1 candidate genes' expression. Generally the F6'H1 candidates expression levels are greater in transgenic cassava root samples with a smaller size, while in larger roots the chance of a lowered F6'H1 candidates expression is greater (figure 5.8). In root slices with diameters greater than 2.5 cm, the F6'H1 candidates mRNA levels were seemingly lower, but the difference was insignificant. However, smaller samples (diameters  $\leq 1.5$  cm) exhibited significantly higher F6'H1 candidate genes' expression levels compared to larger samples or average levels (table 5.3). This variation may be due to the different gene expression behaviours of cassava roots at different maturity stages, and in field-grown mature roots the levels of gene expression may be more consistent.

P values	Wild-type 3DAH	Wild-type 4DAH	Wild-type 5DAH
35S-4	0.0004	0.0011	0.0258
35S-L	0.0001	0.0002	0.0061
PAT-11	0.0012	0.0015	0.0023

**Table 5.2 T-tests reveal the significant down-regulation of the F6'H1 candidates expression in the three transformant lines.** Numbers in the table indicate the P values of the relative quantity data derived from the qRT-PCR comparison between the transgenics and the wild-type at 3, 4, or 5 days after harvest. P-values between 0.05 and 0.01 (95% confidence) are shown in blue, and values less than 0.01 (99% confidence) are shown in red.

P values	All samples	Diameter <=1.5 cm	Diameter >=2.5 cm
Entire population	0.5	0.0021	0.3476
Diameter <=1.5 cm	0.0021	0.5	0.0160
Diameter >=2.5 cm	0.3476	0.0160	0.5

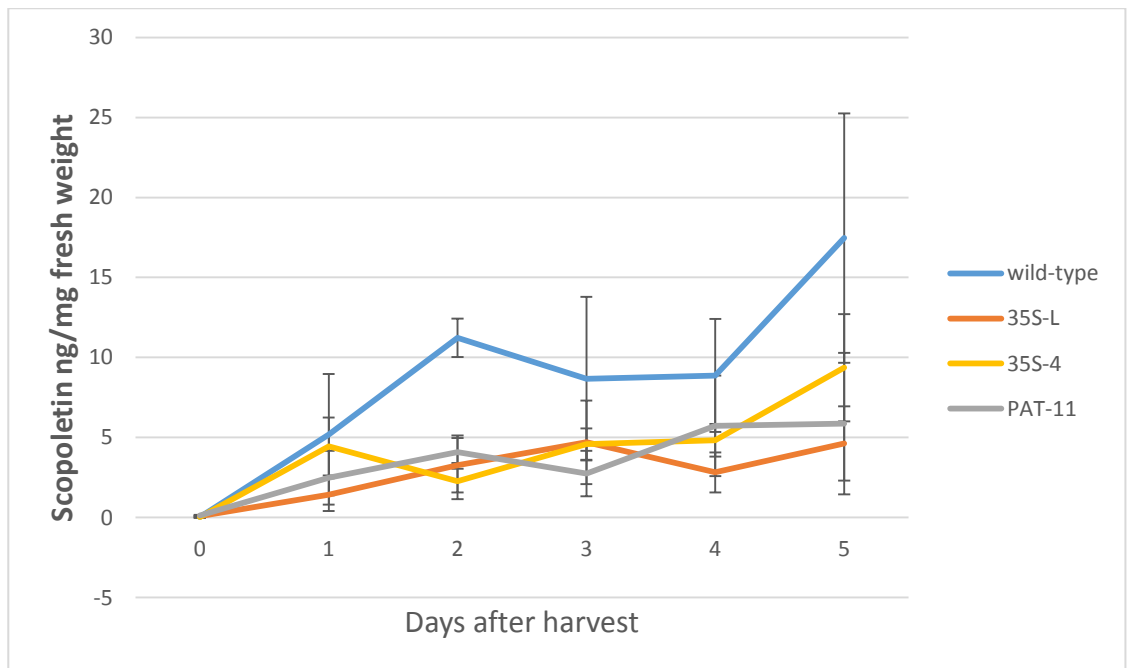
**Table 5.3 T-test revealed the different F6'H1 candidate expression levels in larger (diameters >= 2.5 cm) or smaller (diameters <= 1.5 cm) root samples compared to average.** P-values between 0.05 and 0.01 (95% confidence) are shown in blue, and values less than 0.01 (99% confidence) are shown in red.



**Fig. 5.8 A scatter point plot of the diameter of cassava roots and the expression levels of F6'H1 candidate genes in the RNAi transgenic cassava lines gives a clue on the relation between the size of cassava roots and the expression levels of F6'H1 candidates.** The data points include 62 transgenic cassava root samples from lines 35S-4, 35S-L, and PAT-11 at 3, 4, and 5 days after harvest. (n=62).

### 5.3.5 RNAi down-regulated scopoletin accumulation

The root samples used for the quantitative RT-PCR above were also used for scopoletin quantification. The scopoletin levels in root samples of the wild-type and the three RNAi transgenic lines over the 5 day post-harvest time-course were quantified by LC-MS (figure 5.9).



**Fig. 5.9 The accumulation of scopoletin during the PPD in three selected RNAi mutants and wild-type cassava.** Data are mean scopoletin levels (ng/mg fresh sample)  $\pm$  standard deviation as determined by LC-MS assay from at least four samples ( $n \geq 4$ ).

Despite the great standard deviations, all plants showed a steady increase in scopoletin content over the 5 day time-course from nearly undetectable levels at harvest. Moreover, the transgenic lines containing the RNAi constructs accumulated less scopoletin than wild type, confirming the effectiveness of RNAi in inhibiting scopoletin accumulation (table 5.4). On 2 DAH (the first peak of scopoletin accumulation in cassava roots in figure 5.9), all the three RNAi transgenic lines show significant difference compared to the wild-type. However, during the entire PPD development, only line 35S-L performs significantly down-regulated scopoletin accumulation. Due to time limit and instrument accessibility constraints, only limited numbers of cassava root samples were quantified for their scopoletin accumulation, therefore the limited numbers of replicates may contribute to the large standard deviation and thus the insignificant data. Using other wild-type varieties than TMS 60444 used in these experiments, other researchers have found that scopoletin accumulation peaks around 2 days after harvest and then declines, probably due to exhaustion or tissue damage (Buschmann et al., 2000b, Rickard, 1985). However, it must be pointed out that in

addition using other varieties, the roots used had been grown under different conditions and the preparation of the roots for time-course experiments were also different to those used here. Therefore, while the order of events may be comparable, their details of timing may be different. The apparent transitory decline in scopoletin levels on 3 DAH and 4 DAH in the wild-type is likely to be due to errors in scopoletin quantification considering the large standard deviations of the data (figure 5.9). However, a similar pattern of the declined scopoletin levels was also observed in another research project in the same cultivar (Al-halaseh, unpublished data). Therefore, the reason of such a phenomenon is still unclear.

<b>P values</b>	<b>Wild-type 1DAH</b>	<b>Wild-type 2DAH</b>	<b>Wild-type 3DAH</b>	<b>Wild-type 4DAH</b>	<b>Wild-type 5DAH</b>
<b>35S-4</b>	0.8166	0.0001	0.2790	0.0910	0.1895
<b>35S-L</b>	0.1199	0.0001	0.2025	0.0063	0.0134
<b>PAT-11</b>	0.4096	0.0016	0.0873	0.2589	0.0517

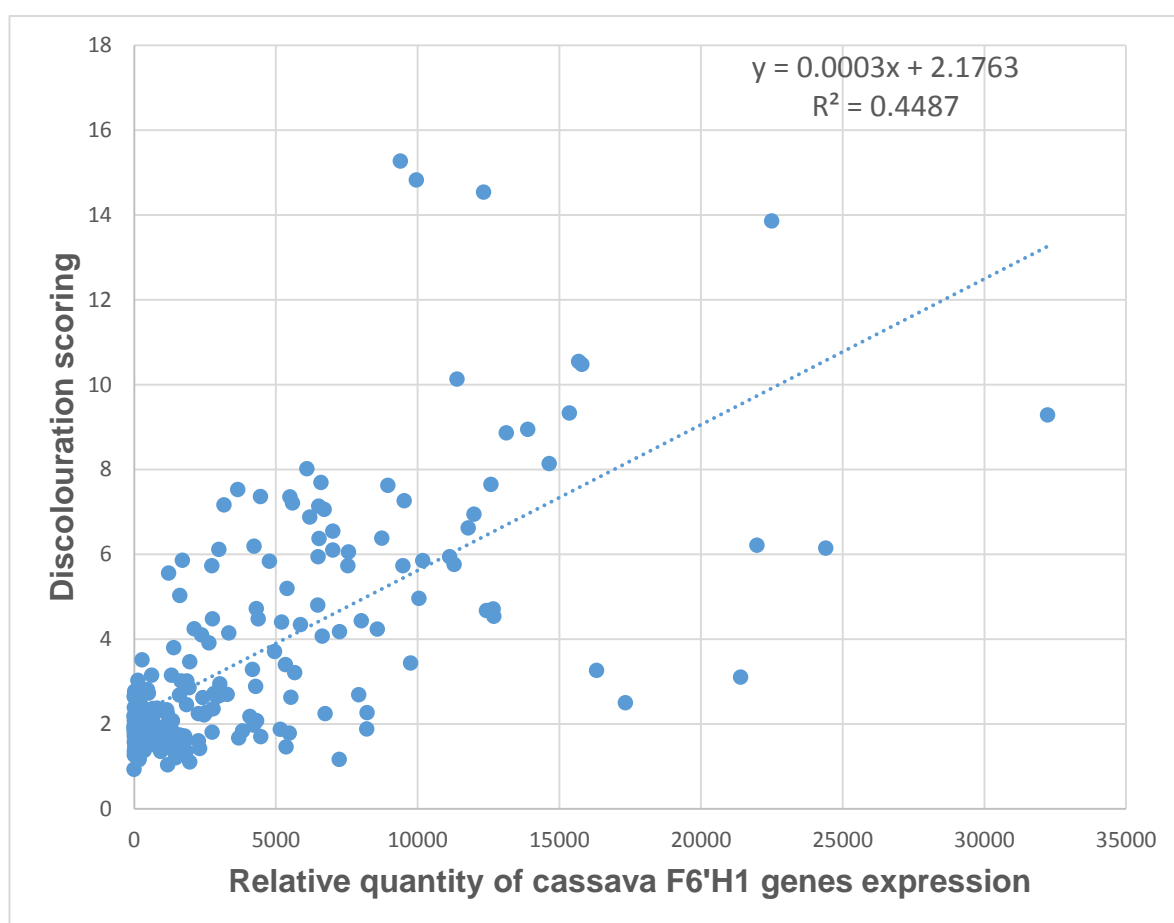
**Table 5.4 T-tests reveal the significant down-regulation of the scopoletin accumulation in the three RNAi transgenic lines.** Numbers in the table indicate the P values of the scopoletin quantification data derived from the LC-MS comparison between the transgenics and the wild-type in a time-course of 5 days after harvest (DAH). P-values between 0.05 and 0.01 are shown in blue and values less than 0.01 are shown in red.

### 5.3.6 PPD development and scopoletin accumulation are related to the expression of cassava F6'H1 gene family

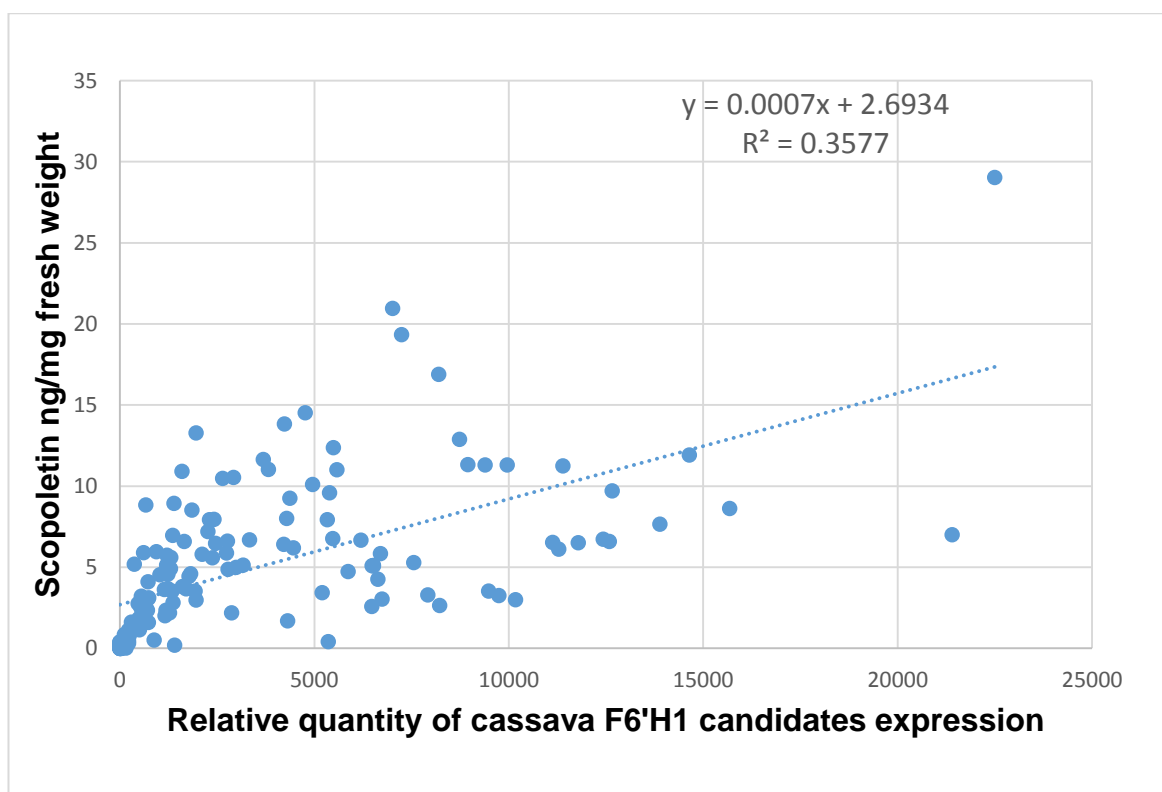
The quantification of PPD discoloration, scopoletin accumulation and cassava F6'H1 genes mRNA levels were carried out on a large number of samples, which provided an opportunity to determine the relationship between the expression of cassava F6'H1 candidates and the development of the other two aspects of PPD development by scatter plots (figure 5.10 and 5.11). In figure 5.10, data from all transgenic lines of cassava (including the six less performing and thus

non-selected lines) were used; while in figure 5.11 only the three chosen lines (i.e. 35S-4, 35S-L, and PAT-11) with the lowest PPD discoloration were used.

As anticipated from the high standard deviations seen in previous figures, the data were noisy. Nonetheless, regression analysis revealed significant trends (P-values less than 0.001, figure 5.10 and 5.11) of correlation between changes in cassava F6'H1 gene expression and the development of discoloration in cassava roots, and between expression and scopoletin accumulation.



**Fig. 5.10 A scatter plot correlates the expression of cassava F6'H1 candidates and the discoloration.** The scopoletin levels (ng/mg fresh sample) and relative abundance of cassava F6'H genes' mRNA was measured from 198 individual samples in LC-MS and qRT-PCR (n=198). Trend line equation:  $y = 0.0003x + 2.1763$ ;  $R^2 = 0.4487$ ; T-stat = 12.63; P-value =  $3.82E-27$ ; Significance F =  $3.82E-27$ .



**Fig. 5.11 A scatter plot correlates the accumulation of scopoletin during the development of PPD and the expression of cassava F6'H1 candidates.** The scopoletin levels (ng/mg fresh sample) and relative abundance of cassava F6'H genes' mRNA was measured from 144 individual samples in LC-MS and qRT-PCR (n=144). Trend line equation:  $y = 0.0007x + 2.6934$ ;  $R^2 = 0.3577$ ; T-stat = 8.89; P-value = 2.47E-15; Significance F = 2.47E-15.

## 5.4 DISCUSSION

This chapter is the culmination of the work described in the preceding chapters, as the preceding chapters discuss preparative work that facilitate the data collecting in this chapter. The development of an RNAi construct targeting most if not all genes in cassava F6'H1 candidate gene family led to the down-regulation of the expression of those genes; however, this down-regulation was not complete. In parallel, in these RNAi transgenic lines, PPD discoloration and scopoletin accumulation were reduced, though not completely prevented.

Putting these data together, the expression of cassava F6'H1 candidates and the accumulation of scopoletin in wild-type cassava roots increase in a time-course of 5 days. A striking discovery was that over the 5 day time-course of PPD development

the expression of the cassava F6'H1 gene family in the wild-type roots increased dramatically from almost undetectable levels on harvest to nearly ten thousand fold, in relative terms, 5 days later. In contrast, the abundance of most of the ROS-detoxifying enzymes increased only 3-5 times during PPD (Owiti et al., 2011). In the RNAi transgenic roots, though the F6'H1 candidates' expression was significantly reduced compared to that in the wild-type, a few thousand fold increase over the background level was still observed. This striking observation may be due to the very low levels of background expression before PPD was triggered. To support this, F6'H1 candidates mRNA abundance on 3 DAH in wild-type is only approximately 3 fold compared to that of 1 DAH, which agrees with the literature (Owiti et al., 2011). In addition, in yeast and human cells high transcription levels of genes do not guarantee proportionally high abundance of the proteins for which they encode (Gygi et al., 1999, Vogel and Marcotte, 2012). Considering the fact that the protein synthetic mechanism is exhaustible, the increase of F6'H1 gene expression should be much less compared to that of the mRNA increase (Bally et al., 2009, Oey et al., 2009). Nonetheless, although the exact abundance of the cassava F6'H1 candidate-encoded proteins is unknown, a significant up-regulation can be predicted from the increase in gene expression. These data indicate that the cassava F6'H1 candidates are expressed at background levels at harvest and are highly PPD-inducible; presumably this increase in mRNA abundance is reflected in parallel accumulation of the corresponding enzymes. An overall conclusion is that cassava F6'H1 gene family is strongly involved in PPD, as knocking the expression of family members down reduced scopoletin accumulation and decreased the PPD response.

On the technical side, the attempt to knock out multiple cassava F6'H1 candidates with a single RNAi construct based on the sequence of one family member showed significant but incomplete effects. Once introduced into the organism, RNAi may target not only its designed target sharing 100 per cent identity to the RNAi transgene, but also other genes with similar sequences (Jackson et al., 2003). Given the fact that plant RNA silencing mechanism permits several mismatches between the small RNA and the target mRNA, the exogenous RNAi may target a larger group of genes (Schwab et al., 2006, Schwab et al., 2005).

The precise mechanism of exogenous dsRNA interference is still unclear, but in *Caenorhabditis elegans* exogenous dsRNA fragments of approximately 500 bp in



length showed the strongest gene silencing, while shorter dsRNA fragments were less efficient at silencing the target gene (Tuschl et al., 1999). This indicates that longer exogenous dsRNA fragments may either produce more small RNA fragments or bind to the target mRNA more firmly. If this is the mechanism, our 471 bp dsRNA should have a considerable chance to target not only gene 10376 but also other genes in the cassava F6'H1 gene family. The siRNA (small interfering RNA) method, which introduces synthetic small dsRNA fragments to target cells to interfere with gene expression, is known to down-regulate multiple genes that share similar sequences in HeLa cells (Jackson et al., 2006). In *Arabidopsis*, an *in silico* study discovered that many genes are able to silence up to 30 related genes if parts of their sequences are used in RNAi (Xu et al., 2006). In our case, the off-target effect of RNAi was more a blessing than a curse, as it had proved challenging to identify with certainty which of the 7 potential candidates were functional F6'H1 orthologs in cassava roots. However, the flip side of this high similarity between the cassava candidate genes is the impossibility to design gene-specific primers that differentiate between the expression profiles of the individual members of the gene family. Therefore, primers targeting the conserved sequence found in all family members were used in the quantitative RT-PCR assays. As a result, the 'broad spectrum' inhibition potential of the RNAi construct helped us study the expression of the cassava F6'H1 gene family holistically.

Despite these challenges, our hypothesis is supported by our data. During cassava PPD development, the cassava F6'H1 candidate gene family is actively involved in the accumulation of scopoletin, which plays an important role in the development of the PPD discoloration.

## 6 GENERAL DISCUSSION

It has long been known that scopoletin accumulates in cassava roots during the PPD response, but the relation between the accumulation of scopoletin and the development of PPD has not been studied in detail previously. A deeper insight into the link between scopoletin and cassava PPD will contribute to the understanding of the PPD process, and may also ultimately lead to extending the shelf-life of cassava roots. Specifically, this research provides an entry point to the study of the relation between the biosynthesis of scopoletin and PPD development in cassava roots. In this research an aim was to attempt to delay PPD by inhibiting scopoletin biosynthesis in cassava. This aim was achieved through the inactivation of cassava F6'H1 candidates that regulate the 6'-hydroxylation of feruloyl CoA, an important step in the biosynthesis of scopoletin. The functional identity of cassava F6'H1 candidates was confirmed by their complementation of *Arabidopsis* F6'H1 T-DNA insert mutants, thereby restoring scopoletin biosynthesis. RNAi down-regulation of the cassava F6'H1 candidate family led to reduced scopoletin accumulation and discoloration in cassava roots during the PPD response. Significant statistical correlation was detected between the expression of cassava F6'H1 candidates, scopoletin accumulation, and the development of PPD symptoms in cassava roots. These results give rise to some interesting issues to discuss.

### 6.1 THE ROLE OF SCOPOLETIN ON A BROADER HORIZON

In this study we set out to explore the relationship between the accumulation of scopoletin and the cassava PPD development. Scopoletin is clearly involved in the PPD development of cassava. Our results show that the accumulation of scopoletin is a significant factor that contributes to the development of discoloration in cassava roots, and that reduced scopoletin accumulation leads to decreased discoloration. Discoloration, the most visible aspect of PPD, is probably due to the oxidation of scopoletin and other related phenolic compounds that possess hydroxyl groups on their aromatic rings in the presence of ROS and peroxidase (Beeching et al., 2002, Chen and Pignatello, 1997, Edwards et al., 1997). Scopoletin is actively accumulated through *de novo* synthesis and/or deglycosylation of its inactive glycoside, scopolin, to discourage microbial and

insect invasion, and it may also contribute an anti-oxidant function in maintaining reactive oxygen species homeostasis (Buschmann et al., 2000a, Giesemann et al., 1986, Goy et al., 1993, Olson and Roseland, 1991, Reilly et al., 2003).

However, scopoletin, and other phenolic compounds, may play wider roles in plants. Scopoletin is a secondary metabolite (hydroxycoumarin) that accumulates as a phytoalexin in plants including cassava, *Arabidopsis*, sunflower (*Helianthus annuus*), tobacco, and many others; its synthesis is triggered by mechanical or entomo-chemical stimulation (Blagbrough et al., 2010, Chong et al., 2002, Kai et al., 2006, Sardari et al., 2000). The antifungal (especially against *Fusarium* spp.) activity of scopoletin and some other coumarin-derivative compounds has been confirmed in several plants (Giesemann et al., 1986, Ojala et al., 2000, Shukla et al., 1999). Their antifungal mechanism is complicated, involving interaction and inhibition of DNA and enzymatic activities, but it is known that substitution of functional groups on the aromatic rings is important to the antifungal bioactivity (Sardari et al., 2000). It has been observed that scopoletin inhibits the germination of a range of fungi (Olson and Roseland, 1991). Further evidence supporting scopoletin's antifungal bioactivity is that, upon fungal infection, scopoletin levels increase in the leaves and upper stems of tobacco, where scopoletin and scopolin normally do not accumulate in the absence of stress (Reuveni and Cohen, 1978). In sunflower, scopoletin accumulates in wild-grown mature individuals but not in individuals grown in a controlled environment suite, indicating that scopoletin synthesis is triggered by stresses encountered by plants in the wild (Gutiérrez-Mellado et al., 1996). Highly concentrated scopoletin also deters insects from feeding on scopoletin-containing plants (Tripathi et al., 2011). Evidence shows that scopoletin induces cell apoptosis in various tumoral and normal cells in mammals by activating the apoptosis-related protease caspase-3 (Kim et al., 2005, Li et al., 2009, Manuele et al., 2006). While plants do not have orthologous caspase genes, they do possess caspase-related mechanisms in their programmed cell death responses (Bonneau et al., 2008). Therefore, although far from definitive, scopoletin may also play a role in plant cell apoptosis as a signalling molecule.

It is still unclear if scopoletin is an active antioxidant, or is just passively oxidised due to its abundant presence. Scopoletin behaves like a double-edged-sword in certain plants. On one hand, it is an important anti-microbial agent, as in *Nicotiana* spp. in response to the application of higher levels of exogenous scopoletin, the

plants are more resistant to fungal pathogens (Goy et al., 1993). On the other hand, scopoletin may do harm to the growth of the plant itself; scopoletin inhibits cell division and root elongation in Timothy-grass (*Phleum pratense*) and wheat (*Triticum* spp.), which indicates that it acts as a signalling molecule or phytotoxin (Avers and Goodwin, 1956, Gutiérrez-Mellado et al., 1996, Shukla et al., 1999). It is even exploited as an allelopathic agent by wild oat (*Avena fatua* L.) to inhibit the growth of potential competitors (Pérez and Ormeño-Núñez, 1991). Considering these diverse roles, this may explain why scopoletin is not accumulated in the absence of stress, as it acts in stress responses and accumulates from either *de novo* synthesis or scopolin deglycosylation when the plant is stressed. Scopoletin may not be an ideal anti-oxidant, especially when efficient and non-toxic alternatives, such as ascorbic acid, are available. In addition, the susceptibility of an *Arabidopsis* ascorbate-deficient mutant to some certain pathogens is diminished comparing to wild-type, which suggests that the anti-microbial activity of a metabolite (e.g. scopoletin) is not necessarily related to its anti-oxidation properties (Barth et al., 2004, Conklin et al., 2000). In anti-oxidation assays scopoletin showed inefficient quenching abilities towards some major reactive oxygen species (e.g.  $\text{OH}^\bullet$  and  $\text{O}_2^{\bullet-}$ ) compared to other phenolic compounds, which renders scopoletin's role as an important general anti-oxidant questionable (Lin et al., 2008, Re et al., 1999, Thuong et al., 2010). In conclusion, although still not definitive, these data imply that scopoletin may not be accumulated principally as an indispensable anti-oxidant.

In addition to the functions mentioned above, several phenolic compounds, including scopoletin and esculetin, act as chelators to improve iron uptake (Clemens and Weber, 2016, Fourcroy et al., 2014). Coumaric compounds are actively secreted through an ATP-binding cassette transporter in the roots of *Arabidopsis* and *Oryza sativa* (rice), and these compounds chelate Fe (III) in the rhizosphere to facilitate iron absorption, although the uptake pathway is still unclear (Bashir et al., 2011, Clemens and Weber, 2016). In *Arabidopsis*, expression of gene F6'H1 is triggered in Fe deficiency (Lan et al., 2011). Moreover, an impaired F6'H1 expression in *Arabidopsis* causes inability to utilise insoluble Fe (III), and externally applied scopoletin or esculetin rescue F6'H1-disabled lines from Fe deficiency (Rodriguez-Celma et al., 2013, Schmid et al., 2014). These data confirm the involvement of scopoletin biosynthesis in iron absorption, and thus a role for scopoletin in Fe uptake. However, unlike esculetin, scopoletin lacks a second

hydroxyl group on its aromatic ring, without which efficient chelating may not occur. Therefore, the precise role of scopoletin in ameliorating Fe deficiency through Fe uptake is still not completely clear.

## **6.2 THE CASSAVA F6'H1 CANDIDATES ARE HIGHLY PPD-INDUCIBLE**

A particular dramatic finding was that, during the PPD time-course, the abundance of cassava F6'H1 mRNAs increased several thousand-fold (figure 5.7). This significant finding was based on the quantification of at least four genes (10291, 10292, 10376, and 27567). It is not uncommon for the expression of a gene to change greatly during different stages of development; for example, in a qRT-PCR study of *Arabidopsis* seeds many genes were expressed several thousand-fold more at the final stages compared to the early stages of seed development (Schmid et al., 2005). In cassava, certain genes are activated during the development of the storage roots and the levels of expression may increase up to 16 thousand-fold (Yang et al., 2011). These dramatic changes in gene expression clearly indicate the importance of cassava F6'H1 candidates in the processes in which they are involved. Data on the changes of expression of other genes in the cassava scopoletin biosynthetic pathways are not yet available. Therefore, it is not possible to say whether or not they also undergo similar changes as those of F6'H1 candidates during the development of PPD.

## **6.3 SOURCE AND SINK OF SCOPOLETIN**

An interesting but yet unclarified question is the source and sink of scopoletin in the PPD response. In plants phenylalanine is synthesised in the chloroplast, where prephenate is transaminated to form L-arogenate, and L-arogenate, which is in turn converted to phenylalanine by arogenate dehydratase (Bickel et al., 1978, Jensen, 1986, Schulze-Siebert et al., 1984). However, in the amyloplast, the starch storage plastid in storage roots, the enzymes that are involved in amino acid metabolism also exist (Balmer et al., 2006). Moreover, linamarin, the predominant cyanogenic compound in cassava, also functions to transport nitrogen from cassava leaves, its site of synthesis, to the storage roots. If the synthesis of linamarin in cassava leaves is inhibited, the development of storage roots is suppressed due to impaired

protein synthesis (Sayre et al., 2011). Therefore, phenylalanine, the 'raw material' for scopoletin *de novo* synthesis, is probably synthesised in cassava roots as long as reduced nitrogen and/or amino acids are transported to and/or available in the storage roots (Mifflin and Lea, 1977).

The scopoletin accumulated during the PPD may derive from two sources, the *de novo* synthesis from phenylalanine, and the deglycosylation of scopolin. Scopolin is supposed to accumulate in the roots gradually in cassava growth as a sink and 'pool' of scopoletin, as it does in *Arabidopsis* roots (Kai et al., 2008, Kai et al., 2006). However, a strong fluorescence under UV light does not occur in fresh cassava roots, but emerges within one day once wounded (figure 1.7). Since both scopoletin and scopolin are fluorescent under UV light (Rickard, 1985), this evidence implies that scopolin is not stored in large quantities in fresh cassava roots. Detached cassava roots do possess a limited amount of phenylalanine (approximately 0.6 ng/mg FW), which renders the hypothesis of the *de novo* synthesis of large amount of scopoletin seems unlikely (Ceballos et al., 2006, Sayre et al., 2011, Yeoh, 1996). However, since phenylalanine can be synthesised in cassava roots from the dehydration of L-arogenate or the transamination of other amino acids, the root should be able to bypass this shortage (Kishore and Shah, 1988, Maeda et al., 2011, Mifflin and Lea, 1977). Preliminary data from another research project in this lab shows that scopolin levels may remain relatively stable or only increase slightly in cassava roots during the PPD response, which favours of the hypothesis of *de novo* synthesis (Al-halaseh, unpublished data). Therefore, *de novo* synthesis may be the major source of scopoletin, or even scopolin, accumulation during the PPD. This hypothesis also explains why the expression of F6'H1 candidate genes was hardly detectable when cassava roots are fresh, but was significantly up-regulated during the PPD (figure 5.7).

To our knowledge studies on the products of hydroxycoumarin oxidation are limited. In the presence of H<sub>2</sub>O<sub>2</sub> and peroxidase, scopoletin is initially oxidised to a blue-black radical intermediate, on which the benzene hydroxyl group is oxidised and carries an unpaired electron (Miller et al., 1975). Subsequently the blue-black intermediate is slowly further oxidised to a water-insoluble yellow complex. This phenomenon is also observed in this study, as freshly cut discoloured cassava roots turn to a yellowish colour once exposed in air for 0.5-1.0 hour. Scopoletin dimers were also confirmed in the oxidative products of scopoletin (Miller et al.,

1975). It is unclear why in cassava roots the oxidative product of scopoletin is predominately in the form of the blue-black intermediate, but limited access to oxygen perhaps contributes to the incomplete oxidation. Nonetheless, further work would be required to understand thoroughly the fate of scopoletin accumulated during the PPD response.

## 6.4 FUTURE WORK

This research project ends here, but the results raise several interesting questions that may lead us further.

The focus of this thesis is on the importance of the 6'-hydroxylation of feruloyl CoA in cassava roots on the accumulation of scopoletin, and on the role of that compound in the development of PPD. However, scopoletin is synthesised through not only one major and, potentially, two minor pathways, but also released through the deglycosylation of scopolin. It would be fascinating to obtain data on the relative flux through these alternatives during both storage root development and their contribution to scopoletin accumulation during PPD. In the research described in this thesis neither the expression of F6'H1 candidates was completely blocked nor the accumulation of scopoletin totally prevented. The residual scopoletin may come from three sources: the alternative scopoletin biosynthetic pathways in the phenylpropanoid metabolism (pathway 1 and 2, figure 1.11); leakage via the incompletely RNAi-blocked cassava F6'H1 candidates; and the deglycosylation of scopolin that had gradually accumulated during the development of cassava storage roots. Frustratingly, the data to hand are not sufficient to resolve fully this question.

The scopoletin biosynthesis in cassava roots involves a network of three possible pathways (figure 1.11). The pathway that we targeted with RNAi is the most important one in scopoletin synthesis contributing more than 90% of *de novo* scopoletin biosynthesis in mature wild-type cassava roots (Bayoumi et al., 2008a). However, when the major pathway is blocked, the behaviour of the phenylpropanoid metabolism may alter due to the feedback from accumulating intermediates. Accumulation of intermediate metabolites may up-regulate the two less important pathways and thus facilitate scopoletin biosynthesis.

Another mechanism yet to be studied is the deglycosylation of storage scopolin in cassava roots. Scopoletin *de novo* synthesis is not the only source of scopoletin accumulation in PPD process, and in *Arabidopsis* roots much higher levels of scopolin are accumulated compared to that of scopoletin (Kai et al., 2008, Kai et al., 2006). As a clue, many phytoanticipins are stored in their inert glycoside form, and hydrolysed by a  $\beta$ -glucosidase to release the bioactive compounds (Morant et al., 2008). In tobacco, mobilisation of root-stored scopolin upon fungal infection has been observed (Reuveni and Cohen, 1978). Therefore, scopolin is probably also hydrolysed by its  $\beta$ -glucosidase to release scopoletin upon stresses in cassava roots. However, what we observed in this study indicates that the accumulation of scopolin in cassava storage roots without stress is very limited (figure 1.7). Further studies on the changed levels of scopolin in cassava storage roots, as well as the role of scopolin- $\beta$ -glucosidase, may help answer the question of the involvement of scopolin in the PPD response.

The use of RNAi to block cassava feruloyl CoA 6'-hydroxylation failed to completely knock-out the activity of the cassava F6'H1 genes. Other gene silencing techniques, such as CRISPR-Cas-9, might succeed in silencing the cassava F6'H1 genes completely, thereby providing a clearer insight into the relation between scopoletin synthesis and the discoloration (Barrangou et al., 2007, Cong et al., 2013). In addition, the experimental cassava roots were grown in controlled environments under the temporal and spatial limitations imposed by the glasshouse, and it is possible that the small roots may behave differently in terms of their post-harvest responses than field-grown ones. Field trials allowing transgenic cassava lines to develop naturally might give clearer insights into PPD process.

Scopoletin and many other phenolic compounds accumulate in plants for defensive and other purposes; therefore, blocking their synthesis may impair defence or aspects of development (Bayoumi et al., 2008b, Tomás-Barberán and Espín, 2001). For example, depletion of certain stress-response phenolics may render the plant vulnerable (Winkel-Shirley, 2002b). In *Arabidopsis* mutants in either chalcone synthase or chalcone isomerase prevented the synthesis of flavonols, rendering the mutant lines more vulnerable to UV irradiation (Landry et al., 1995). We also observed altered phenotypes in some F6'H1 candidate RNAi transgenic cassava (figure 5.3). Referring to the broad spectrum of scopoletin's potential bioactivities, blockage of scopoletin biosynthesis may lead to the loss of stress resistance,



including but not limited to oxidative stress, microbial stress and Fe deficiency. On the other hand, altered phenotypes due to reduced scopoletin could be a useful tool for identifying the roles scopoletin plays in the life of plants. The literature on reactive oxygen species (ROS) metabolism in plants with impaired phenolic synthesis is limited, but whether or not down-regulated scopoletin accumulation in cassava roots worsens the oxidative damage and ROS-induced apoptosis, is an interesting question to address in the future (Simon et al., 2000, Van Breusegem and Dat, 2006). In addition, microbial resistance, ferric uptake, and even the flavour of cassava may also be affected by an altered scopoletin and phenolic composition (Clemens and Weber, 2016, Gnonlonfin et al., 2012, Tomás-Barberán and Espín, 2001). Studies on these aspects may lead to further insights in the future.

## 7 REFERENCES

- ABHARY, M., SIRITUNGA, D., STEVENS, G., TAYLOR, N. J. & FAUQUET, C. M. 2011. Transgenic biofortification of the starchy staple cassava (*Manihot esculenta*) generates a novel sink for protein. *PLoS One*, 6, e16256.
- ADELEKAN, B. A. 2010. Investigation of ethanol productivity of cassava crop as a sustainable source of biofuel in tropical countries. *African Journal of Biotechnology*, 9, 5643-5650.
- AGBOR-EGBE, T. O. M. & RICKARD, J. E. 1991. Study on the factors affecting storage of edible aroids. *Annals of Applied Biology*, 119, 121-130.
- AHN, Y. O., SHIMIZU, B., SAKATA, K., GANTULGA, D., ZHOU, C., BEVAN, D. R. & ESEN, A. 2010. Scopolin-hydrolyzing  $\beta$ -glucosidases in roots of *Arabidopsis*. *Plant Cell Physiology*, 51, 132-43.
- ALICAI, T., OMONGO, C. A., MARUTHI, M. N., HILLOCKS, R. J., BAGUMA, Y., KAWUKI, R., BUA, A., OTIM-NAPE, G. W. & COLVIN, J. 2007. Re-emergence of cassava brown streak disease in Uganda. *Plant Disease*, 91, 24-29.
- ALLEM, C. A. 1994. The origin of *Manihot esculenta* Crantz (Euphorbiaceae). *Genetic Resources and Crop Evolution*, 41, 133-150.
- ALONSO, J. M., STEPANOVA, A. N., LEISSE, T. J., KIM, C. J., CHEN, H., SHINN, P., STEVENSON, D. K., ZIMMERMAN, J., BARAJAS, P., CHEUK, R., GADRINAB, C., HELLER, C., JESKE, A., KOESEMA, E., MEYERS, C. C., PARKER, H., PREDNIS, L., ANSARI, Y., CHOY, N., DEEN, H., GERALT, M., HAZARI, N., HOM, E., KARNES, M., MULHOLLAND, C., NDUBAKU, R., SCHMIDT, I., GUZMAN, P., AGUILAR-HENONIN, L., SCHMID, M., WEIGEL, D., CARTER, D. E., MARCHAND, T., RISSEUW, E., BROGDEN, D., ZEKO, A., CROSBY, W. L., BERRY, C. C. & ECKER, J. R. 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*, 301, 653-7.
- ALOYS, N. & MING, Z. H. 2006. Traditional cassava foods in Burundi - A review. *Food Reviews International*, 22, 1-27.
- ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *Journal of Molecular Biology*, 215, 403-10.
- ALVES, A. A. C. 2002. Cassava Botany and Physiology. In: HILLOCKS, R. J., THRESH, J. M. & BELLOTTI, A. (eds.) *Cassava : biology, production and utilization*. Wallingford: CABI Pub.
- AVERS, C. J. & GOODWIN, R. H. 1956. Studies on roots. IV. Effects of coumarin and scopoletin on the standard root growth pattern of *Phleum pratense*. *American Journal of Botany*, 43, 612-620.
- BALASUNDRAM, N., SUNDRAM, K. & SAMMAN, S. 2006. Phenolic compounds in plants and agri-industrial by-products: Antioxidant activity, occurrence, and potential uses. *Food Chemistry*, 99, 191-203.
- BALLY, J., NADAI, M., VITEL, M., ROLLAND, A., DUMAIN, R. & DUBALD, M. 2009. Plant physiological adaptations to the massive foreign protein synthesis occurring in recombinant chloroplasts. *Plant Physiology*, 150, 1474-81.
- BALMER, Y., VENSEL, W. H., CAI, N., MANIERI, W., SCHURMANN, P., HURKMAN, W. J. & BUCHANAN, B. B. 2006. A complete ferredoxin/thioredoxin system regulates fundamental processes in amyloplasts. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 2988-93.
- BARRANGOU, R., FREMAUX, C., DEVEAU, H., RICHARDS, M., BOYAVAL, P., MOINEAU, S., ROMERO, D. A. & HORVATH, P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 315, 1709-12.
- BARRE, W. L. 1938. Native american beers. *American Anthropologist*, 40, 224-234.
- BARTH, C., MOEDER, W., KLESSIG, D. F. & CONKLIN, P. L. 2004. The timing of senescence and response to pathogens is altered in the ascorbate-deficient *Arabidopsis* mutant vitamin c-1. *Plant Physiology*, 134, 1784-92.

- BASHIR, K., ISHIMARU, Y., SHIMO, H., KAKEI, Y., SENOURA, T., TAKAHASHI, R., SATO, Y., SATO, Y., UOZUMI, N., NAKANISHI, H. & NISHIZAWA, N. K. 2011. Rice phenolics efflux transporter 2 (PEZ2) plays an important role in solubilizing apoplasmic iron. *Soil Science and Plant Nutrition*, 57, 803-812.
- BAYOUMI, S. A., ROWAN, M. G., BEECHING, J. R. & BLAGBROUGH, I. S. 2008a. Investigation of biosynthetic pathways to hydroxycoumarins during post-harvest physiological deterioration in Cassava roots by using stable isotope labelling. *ChemBioChem*, 9, 3013-22.
- BAYOUMI, S. A., ROWAN, M. G., BEECHING, J. R. & BLAGBROUGH, I. S. 2010. Constituents and secondary metabolite natural products in fresh and deteriorated cassava roots. *Phytochemistry*, 71, 598-604.
- BAYOUMI, S. A., ROWAN, M. G., BLAGBROUGH, I. S. & BEECHING, J. R. 2008b. Biosynthesis of scopoletin and scopolin in cassava roots during post-harvest physiological deterioration: the E-Z-isomerisation stage. *Phytochemistry*, 69, 2928-36.
- BEECHING, J. R., HAN, Y. H., GOMEZ-VASQUEZ, R., DAY, R. C. & COOPER, R. M. 1998. Wound and defense responses in cassava as related to post-harvest physiological deterioration. *Phytochemical Signals and Plant-Microbe Interactions*, 32, 231-248.
- BEECHING, J. R., REILLY, K., GOMEZ-VASQUEZ, R., LI, H., HAN, Y., RODRIGUEZ, M. X., BUSCHMANN, H., TAYLOR, N., FAUQUET, C. & TOHME, J. Post-harvest physiological deterioration of cassava. 12th Symposium of the International Society for Tropical Root Crops: Potential of root crops for food and industrial resources (10–12 September 2000, Tsukuba, Japan), Cultio, Tsukuba, 2002. 60-66.
- BELL-LELONG, D. A., CUSUMANO, J. C., MEYER, K. & CHAPPLE, C. 1997. Cinnamate-4-hydroxylase expression in *Arabidopsis*. Regulation in response to development and the environment. *Plant Physiology*, 113, 729-38.
- BENITO, I. R., OMAR, G. O. & CARLO, S. R. 2013. Characterization of dietary fiber and pectin of cassava bread obtained from different regions of Venezuela. *Revista chilena de nutrición*, 40, 169-173.
- BICKEL, H., PALME, L. & SCHULTZ, G. 1978. Incorporation of shikimate and other precursors into aromatic amino acids and prenylquinones of isolated spinach chloroplasts. *Phytochemistry*, 17, 119-124.
- BIOMATTERS 2016. Geneious 9.1 Manual.
- BLAGBROUGH, I. S., BAYOUMI, S. A., ROWAN, M. G. & BEECHING, J. R. 2010. Cassava: an appraisal of its phytochemistry and its biotechnological prospects. *Phytochemistry*, 71, 1940-51.
- BONNEAU, L., GE, Y., DRURY, G. E. & GALLOIS, P. 2008. What happened to plant caspases? *Journal of Experimental Botany*, 59, 491-9.
- BOX, J. F. 1987. Guinness, Gosset, Fisher, and small samples. *Statistical Science*, 2, 45-52.
- BULL, S. E. 2011. *Study of post-harvest physiological deterioration in transgenic cassava*. PhD, University of Bath.
- BULL, S. E., NDUNGURU, J., GRUISSEM, W., BEECHING, J. R. & VANDERSCHUREN, H. 2011. Cassava: constraints to production and the transfer of biotechnology to African laboratories. *Plant Cell Reports*, 30, 779-87.
- BULL, S. E., OWITI, J. A., NIKLAUS, M., BEECHING, J. R., GRUISSEM, W. & VANDERSCHUREN, H. 2009. *Agrobacterium*-mediated transformation of friable embryogenic calli and regeneration of transgenic cassava. *Nature Protocols*, 4, 1845-54.
- BUSCHMANN, H., REILLY, K., RODRIGUEZ, M. X., TOHME, J. & BEECHING, J. R. 2000a. Hydrogen peroxide and flavan-3-ols in storage roots of cassava (*Manihot esculenta* Crantz) during postharvest deterioration. *Journal of Agricultural and Food Chemistry*, 48, 5522-9.
- BUSCHMANN, H., RODRIGUEZ, M. X., TOHME, J. & BEECHING, J. R. 2000b. Accumulation of hydroxycoumarins during post-harvest deterioration of tuberous roots of Cassava (*Manihot esculenta* Crantz). *Annals of Botany*, 86, 1153-1160.
- CAI, Y., LUO, Q., SUN, M. & CORKE, H. 2004. Antioxidant activity and phenolic compounds of 112 traditional Chinese medicinal plants associated with anticancer. *Life Sciences*, 74, 2157-84.

- CARDOSO, A. P., MIRIONE, E., ERNESTO, M., MASSAZA, F., CLIFF, J., HAQUE, M. R. & BRADBURY, J. H. 2005. Processing of cassava roots to remove cyanogens. *Journal of Food Composition and Analysis*, 18, 451-460.
- CASS, C. L., PERALDI, A., DOWD, P. F., MOTTIAR, Y., SANTORO, N., KARLEN, S. D., BUKHMAN, Y. V., FOSTER, C. E., THROWER, N., BRUNO, L. C., MOSKVIN, O. V., JOHNSON, E. T., WILLHOIT, M. E., PHUTANE, M., RALPH, J., MANSFIELD, S. D., NICHOLSON, P. & SEDBROOK, J. C. 2015. Effects of phenylalanine ammonia lyase (PAL) knockdown on cell wall composition, biomass digestibility, and biotic and abiotic stress responses in *Brachypodium*. *Journal of Experimental Botany*, 66, 4317-35.
- CEBALLOS, H., IGLESIAS, C. A., PEREZ, J. C. & DIXON, A. G. O. 2004. Cassava breeding: opportunities and challenges. *Plant Molecular Biology*, 56, 503-516.
- CEBALLOS, H., SANCHEZ, T., CHAVEZ, A. L., IGLESIAS, C., DEBOUCK, D., MAFLA, G. & TOHME, J. 2006. Variation in crude protein content in cassava (*Manihot esculenta* Crantz) roots. *Journal of Food Composition and Analysis*, 19, 589-593.
- CHANG, S., PURYEAR, J. & CAIRNEY, J. 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter*, 11, 113-116.
- CHAVEZ, A. L., SANCHEZ, T., JARAMILLO, G., BEDOYA, J. M., ECHEVERRY, J., BOLANOS, E. A., CEBALLOS, H. & IGLESIAS, C. A. 2005. Variation of quality traits in cassava roots evaluated in landraces and improved clones. *Euphytica*, 143, 125-133.
- CHEN, R. & PIGNATELLO, J. J. 1997. Role of quinone intermediates as electron shuttles in Fenton and photoassisted Fenton oxidations of aromatic compounds. *Environmental Science & Technology*, 31, 2399-2406.
- CHONG, J., BALTZ, R., SCHMITT, C., BEFFA, R., FRITIG, B. & SAINDRENAN, P. 2002. Downregulation of a pathogen-responsive tobacco UDP-Glc: Phenylpropanoid glucosyltransferase reduces scopoletin glucoside accumulation, enhances oxidative stress, and weakens virus resistance. *The Plant Cell*, 14, 1093-1107.
- CLEMENS, S. & WEBER, M. 2016. The essential role of coumarin secretion for Fe acquisition from alkaline soil. *Plant Signaling & Behavior*, 11, e1114197.
- CLEMENT, C. R., DE CRISTO-ARA JO, M., D'EECKENBRUGGE, G. C., ALVES PEREIRA, A. & PICAN O-RODRIGUES, D. 2010. Origin and Domestication of Native Amazonian Crops. *Diversity*, 2, 72-106.
- CLIFF, J., MUQUINGUE, H., NHASSICO, D., NZWALO, H. & BRADBURY, J. H. 2011. Konzo and continuing cyanide intoxication from cassava in Mozambique. *Food and Chemical Toxicology*, 49, 631-5.
- CLOUGH, S. J. & BENT, A. F. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal*, 16, 735-43.
- COCHRANE, F. C., DAVIN, L. B. & LEWIS, N. G. 2004. The *Arabidopsis* phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. *Phytochemistry*, 65, 1557-64.
- COCK, J. H. 1982. Cassava: a basic energy source in the tropics. *Science*, 218, 755-62.
- COLEMAN, H. D., PARK, J. Y., NAIR, R., CHAPPLE, C. & MANSFIELD, S. D. 2008. RNAi-mediated suppression of p-coumaroyl-CoA 3'-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 4501-6.
- CONG, L., RAN, F. A., COX, D., LIN, S., BARRETTO, R., HABIB, N., HSU, P. D., WU, X., JIANG, W., MARRAFFINI, L. A. & ZHANG, F. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339, 819-23.
- CONKLIN, P. L., SARACCO, S. A., NORRIS, S. R. & LAST, R. L. 2000. Identification of ascorbic acid-deficient *Arabidopsis thaliana* mutants. *Genetics*, 154, 847-56.
- COOKE, R. D., BLAKE, G. G. & BATTERSHILL, J. M. 1978. Purification of cassava linamarase. *Phytochemistry*, 17, 381-383.
- COOPER, C. E. & BROWN, G. C. 2008. The inhibition of mitochondrial cytochrome oxidase by the gases carbon monoxide, nitric oxide, hydrogen cyanide and hydrogen sulfide: chemical

- mechanism and physiological significance. *Journal of Bioenergetics and Biomembranes*, 40, 533-9.
- COTTLE, W. & KOLATTUKUDY, P. E. 1982. Absciscic acid stimulation of suberization : induction of enzymes and deposition of polymeric components and associated waxes in tissue cultures of potato tuber. *Plant Physiology*, 70, 775-80.
- COWAN, M. M. 1999. Plant products as antimicrobial agents. *Clinical Microbiology Reviews*, 12, 564-582.
- CZECHOWSKI, T., STITT, M., ALTMANN, T., UDVARDI, M. K. & SCHEIBLE, W. R. 2005. Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiology*, 139, 5-17.
- DE BOLLE, M. F. C., BUTAYE, K. M. J., COUCKE, W. J. W., GODERIS, I. J. W. M., WOUTERS, P. F. J., VAN BOXEL, N., BROEKAERT, W. F. & CAMMUE, B. P. A. 2003. Analysis of the influence of promoter elements and a matrix attachment region on the inter-individual variation of transgene expression in populations of *Arabidopsis thaliana*. *Plant Science*, 165, 169-179.
- DEKKERS, B. J., WILLEMS, L., BASSEL, G. W., VAN BOLDEREN-VELDKAMP, R. P., LIGTERINK, W., HILHORST, H. W. & BENTSINK, L. 2012. Identification of reference genes for RT-qPCR expression analysis in *Arabidopsis* and tomato seeds. *Plant Cell Physiology*, 53, 28-37.
- DOERKS, T., COPLEY, R. R., SCHULTZ, J., PONTING, C. P. & BORK, P. 2002. Systematic identification of novel protein domain families associated with nuclear functions. *Genome Research*, 12, 47-56.
- DOREY, S. 1999. Hydrogen peroxide from the oxidative burst is neither necessary nor sufficient for hypersensitive cell death induction, phenylalanine ammonia lyase stimulation, salicylic acid accumulation, or scopoletin consumption in cultured tobacco cells treated with elicitor. *Plant Physiology*, 121, 163-172.
- DUFOUR, D. L. "Bitter" cassava: Toxicity and detoxification. In: NASSAR, N. & ORTIZ, R., eds. First International Meeting on Cassava Breeding, Biotechnology and Ecology, 2006 Brasilia. Thesaurus Editora, 171-184.
- EDWARDS, R., STONES, S. M., GUTIERREZMELLADO, M. C. & JORRIN, J. 1997. Characterization and inducibility of a scopoletin-degrading enzyme from sunflower. *Phytochemistry*, 45, 1109-1114.
- EGGUM, B. O. 1970. The protein quality of cassava leaves. *British Journal of Nutrition*, 24, 761-8.
- EJECHI, B. O. & SOUZEY, J. A. 1999. Inhibition of biodeterioration of yam tuber *Dioscorea rotundata* Poir in storage with phenolic extract of *Acalypha hispida* Burm.f. leaves. *Journal of Stored Products Research*, 35, 127-134.
- EL-SHARKAWY, M. A. 2004. Cassava biology and physiology. *Plant Molecular Biology*, 56, 481-501.
- EL-SHARKAWY, M. A. 2006. International research on cassava photosynthesis, productivity, eco-physiology, and responses to environmental stresses in the tropics. *Photosynthetica*, 44, 481-512.
- ELBASHIR, S. M., LENDECKEL, W. & TUSCHL, T. 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & Development*, 15, 188-200.
- EMILIANI, G., FONDI, M., FANI, R. & GRIBALDO, S. 2009. A horizontal gene transfer at the origin of phenylpropanoid metabolism: a key adaptation of plants to land. *Biology Direct*, 4, 7.
- ERNESTO, M., CARDOSO, A. P., NICALA, D., MIRIONE, E., MASSAZA, F., CLIFF, J., HAQUE, M. R. & BRADBURY, J. H. 2002. Persistent konzo and cyanogen toxicity from cassava in northern Mozambique. *Acta Tropica*, 82, 357-62.
- FAO. 2008. *International Year of the Potato* [Online]. Available: <http://www.fao.org/potato-2008/en/world/> [Accessed 11/Nov 2015].
- FAOSTAT. 2013. *Cassava Production share by region* [Online]. Available: <http://faostat3.fao.org/browse/Q/QC/E> [Accessed 16/11 2015].
- FERRER, J. L., AUSTIN, M. B., STEWART, C., JR. & NOEL, J. P. 2008. Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiology and Biochemistry*, 46, 356-70.

- FIRE, A., XU, S., MONTGOMERY, M. K., KOSTAS, S. A., DRIVER, S. E. & MELLO, C. C. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, 391, 806-11.
- FOURCROY, P., SISO-TERRAZA, P., SUDRE, D., SAVIRON, M., REYT, G., GAYMARD, F., ABADIA, A., ABADIA, J., ALVAREZ-FERNANDEZ, A. & BRIAT, J. F. 2014. Involvement of the ABCG37 transporter in secretion of scopoletin and derivatives by *Arabidopsis* roots in response to iron deficiency. *New Phytologist*, 201, 155-67.
- FRANKE, R. & SCHREIBER, L. 2007. Suberin--a biopolyester forming apoplastic plant interfaces. *Current Opinion in Plant Biology*, 10, 252-9.
- FRITIG, B., HIRTH, L. & OURISSON, G. 1970. Biosynthesis of the coumarins: Scopoletin formation in tobacco tissue cultures. *Phytochemistry*, 9, 1963-1975.
- GAZZANI, S., LAWRENSEN, T., WOODWARD, C., HEADON, D. & SABLOWSKI, R. 2004. A link between mRNA turnover and RNA interference in *Arabidopsis*. *Science*, 306, 1046-8.
- GERTZ, E. M., YU, Y. K., AGARWALA, R., SCHAFFER, A. A. & ALTSCHUL, S. F. 2006. Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST. *BMC Biology*, 4, 41.
- GIESEMANN, A., BIEHL, B. & LIEBEREI, R. 1986. Identification of scopoletin as a phytoalexin of the rubber tree *Hevea brasiliensis*. *Journal of Phytopathology*, 117, 373-376.
- GNONLONFIN, B. G., GBAGUIDI, F., GBENOU, J. D., SANNNI, A. & BRIMER, L. 2011. Changes in scopoletin concentration in cassava chips from four varieties during storage. *Journal of the Science of Food and Agriculture*, 91, 2344-7.
- GNONLONFIN, B. G. J., SANNNI, A. & BRIMER, L. 2012. Review scopoletin - a coumarin phytoalexin with medicinal properties. *Critical Reviews in Plant Sciences*, 31, 47-56.
- GOMEZ-VASQUEZ, R., DAY, R., BUSCHMANN, H., RANGLES, S., BEECHING, J. R. & COOPER, R. M. 2004. Phenylpropanoids, Phenylalanine Ammonia Lyase and Peroxidases in Elicitor-challenged Cassava (*Manihot esculenta*) Suspension Cells and Leaves. *Annals of Botany*, 94, 87-97.
- GOODSTEIN, D. M., SHU, S., HOWSON, R., NEUPANE, R., HAYES, R. D., FAZO, J., MITROS, T., DIRKS, W., HELLSTEN, U., PUTNAM, N. & ROKHSAR, D. S. 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, 40, D1178-86.
- GOUY, M., GUINDON, S. & GASCUEL, O. 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, 27, 221-4.
- GOY, P. A., SIGNER, H., REIST, R., AICHHOLZ, R., BLUM, W., SCHMIDT, E. & KESSMANN, H. 1993. Accumulation of scopoletin is associated with the high disease resistance of the hybrid *Nicotiana glutinosa* x *Nicotiana debneyi*. *Planta*, 191.
- GUTIERREZ-MELLADO, M.-C., EDWARDS, R., TENA, M., CABELLO, F., SERGHINI, K. & JORR N, J. 1996. The production of coumarin phytoalexins in different plant organs of sunflower (*Helianthus annuus* L.). *Journal of Plant Physiology*, 149, 261-266.
- GUTIERREZ, M. 1995. Abiotic elicitation of coumarin phytoalexins in sunflower. *Phytochemistry*, 38, 1185-1191.
- GYGI, S. P., ROCHON, Y., FRANZA, B. R. & AEBSOLD, R. 1999. Correlation between protein and mRNA abundance in yeast. *Molecular and Cellular Biology*, 19, 1720-1730.
- HAHLBROCK, K. & SCHEEL, D. 1989. Physiology and Molecular Biology of Phenylpropanoid Metabolism. *Annual Review of Plant Physiology and Plant Molecular Biology*, 40, 347-369.
- HAMILTON, A. J. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, 286, 950-952.
- HASEGAWA, M. & FUJIWARA, M. 1993. Relative efficiencies of the maximum likelihood, maximum parsimony, and neighbor-joining methods for estimating protein phylogeny. *Molecular Phylogenetics and Evolution*, 2, 1-5.
- HELLIWELL, C. & WATERHOUSE, P. M. 2003. Constructs and methods for high-throughput gene silencing in plants. *Methods*, 30, 289-295.

- HIEI, Y., OHTA, S., KOMARI, T. & KUMASHIRO, T. 1994. Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *The Plant Journal*, 6, 271-82.
- HILLOCKS, R. J. 2002. Cassava in Africa. In: HILLOCKS, R. J., THRESH, J. M. & BELLOTTI, A. (eds.) *Cassava : biology, production and utilization*. Wallingford: CABI Pub.
- HILLOCKS, R. J. & JENNINGS, D. L. 2003. Cassava brown streak disease: a review of present knowledge and research needs. *International Journal of Pest Management*, 49, 225-234.
- HIROSE, S. 1986. Physiological studies on post-harvest deterioration of cassava roots. *Japan Agricultural Research Quarterly*, 19, 241-252.
- HOBBS, S. L. A., WARKENTIN, T. D. & DELONG, C. M. O. 1993. Transgene copy number can be positively or negatively associated with transgene expression. *Plant Molecular Biology*, 21, 17-26.
- HOEKEMA, A., HIRSCH, P. R., HOOYKAAS, P. J. J. & SCHILPEROORT, R. A. 1983. A binary plant vector strategy based on separation of vir- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature*, 303, 179-180.
- HOLMES, G. J. & STANGE, R. R. 2002. Influence of wound type and storage duration on susceptibility of sweetpotatoes to *Rhizopus* soft rot. *Plant Disease*, 86, 345-348.
- HUMPHREYS, J. M. & CHAPPLE, C. 2002. Rewriting the lignin roadmap. *Current Opinion in Plant Biology*, 5, 224-9.
- HUNT, L. A., WHOLEY, D. W. & COCK, J. H. 1977. Growth physiology of cassava (*Manihot esculenta* Crantz). *Field Crop Abstracts*, 30, 77-91.
- IKEDIABI, C. O., CHELVARAJAN, R. L. & UKOHA, A. I. 1989. Biochemical aspects of wound-healing in yams (*Dioscorea* spp.). *Journal of the Science of Food and Agriculture*, 48, 131-139.
- IMASEKI, H. 1970. Induction of peroxidase activity by ethylene in sweet potato. *Plant Physiology*, 46, 172-4.
- IMASEKI, H., URITANI, I. & STAHMANN, M. A. 1968. Production of ethylene by injured sweet potato root tissue. *Plant and Cell Physiology*, 9, 757-&.
- JACKSON, A. L., BARTZ, S. R., SCHELTER, J., KOBAYASHI, S. V., BURCHARD, J., MAO, M., LI, B., CAVET, G. & LINSLEY, P. S. 2003. Expression profiling reveals off-target gene regulation by RNAi. *Nature Biotechnology*, 21, 635-7.
- JACKSON, A. L., BURCHARD, J., SCHELTER, J., CHAU, B. N., CLEARY, M., LIM, L. & LINSLEY, P. S. 2006. Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA*, 12, 1179-87.
- JANSSON, C., WESTERBERGH, A., ZHANG, J. M., HU, X. W. & SUN, C. X. 2009. Cassava, a potential biofuel crop in (the) People's Republic of China. *Applied Energy*, 86, S95-S99.
- JEFFERSON, R., GOLDSBROUGH, A. & BEVAN, M. 1990. Transcriptional regulation of a patatin-1 gene in potato. *Plant Molecular Biology*, 14, 995-1006.
- JEKAYINFA, S. O. & OLAJIDE, J. O. 2007. Analysis of energy usage in the production of three selected cassava-based foods in Nigeria. *Journal of Food Engineering*, 82, 217-226.
- JENSEN, R. A. 1986. The shikimate/arogenate pathway: Link between carbohydrate metabolism and secondary metabolism. *Physiologia Plantarum*, 66, 164-168.
- JORGENSEN, R. 1990. Altered gene expression in plants due totans interactions between homologous genes. *Trends in Biotechnology*, 8, 340-344.
- KAI, K., MIZUTANI, M., KAWAMURA, N., YAMAMOTO, R., TAMAI, M., YAMAGUCHI, H., SAKATA, K. & SHIMIZU, B. 2008. Scopoletin is biosynthesized via ortho-hydroxylation of feruloyl CoA by a 2-oxoglutarate-dependent dioxygenase in *Arabidopsis thaliana*. *The Plant Journal*, 55, 989-99.
- KAI, K., SHIMIZU, B., MIZUTANI, M., WATANABE, K. & SAKATA, K. 2006. Accumulation of coumarins in *Arabidopsis thaliana*. *Phytochemistry*, 67, 379-86.
- KALIDAS, S., LI, Q. & PHILLIPS, M. A. 2011. A Gateway® compatible vector for gene silencing in bloodstream form *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, 178, 51-5.
- KARIMI, M., INZE, D. & DEPICKER, A. 2002. GATEWAY™ vectors for *Agrobacterium*-mediated plant transformation. *Trends in Plant Science*, 7, 193-5.

- KAY, R., CHAN, A., DALY, M. & MCPHERSON, J. 1987. Duplication of CaMV 35S Promoter Sequences Creates a Strong Enhancer for Plant Genes. *Science*, 236, 1299-302.
- KEARSE, M., MOIR, R., WILSON, A., STONES-HAVAS, S., CHEUNG, M., STURROCK, S., BUXTON, S., COOPER, A., MARKOWITZ, S., DURAN, C., THIERER, T., ASHTON, B., MEINTJES, P. & DRUMMOND, A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28, 1647-9.
- KIM, E. K., KWON, K. B., SHIN, B. C., SEO, E. A., LEE, Y. R., KIM, J. S., PARK, J. W., PARK, B. H. & RYU, D. G. 2005. Scopoletin induces apoptosis in human promyeloleukemic cells, accompanied by activations of nuclear factor kappaB and caspase-3. *Life Sciences*, 77, 824-36.
- KISHORE, G. M. & SHAH, D. M. 1988. Amino acid biosynthesis inhibitors as herbicides. *Annual Review of Biochemistry*, 57, 627-63.
- KLAHRE, U., CRETE, P., LEUENBERGER, S. A., IGLESIAS, V. A. & MEINS, F., JR. 2002. High molecular weight RNAs and small interfering RNAs induce systemic posttranscriptional gene silencing in plants. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 11981-6.
- KONCZ, C. & SCHELL, J. 1986. The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Molecular and General Genetics*, 204, 383-396.
- KUNTTU, L., LEPIST, L., RAUHAMAA, J. & VISA, A. 2003. Binary histogram in image classification for retrieval purposes. *Journal of WSCG*, 11, 269-273.
- LALAGUNA, F. & AGUDO, M. 1989. Relationship between changes in lipid with ageing of cassava roots and senescence parameters. *Phytochemistry*, 28, 2059-2062.
- LAN, P., LI, W., WEN, T. N., SHIAU, J. Y., WU, Y. C., LIN, W. & SCHMIDT, W. 2011. iTRAQ protein profile analysis of *Arabidopsis* roots reveals new aspects critical for iron homeostasis. *Plant Physiology*, 155, 821-34.
- LANCASTER, P. A., INGRAM, J. S., LIM, M. Y. & COURSEY, D. G. 1982. Traditional cassava-based foods: Survey of processing techniques. *Economic Botany*, 36, 12-45.
- LANDRY, L. C., CHAPPLE, C. C. S. & LAST, R. L. 1995. *Arabidopsis* mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-b Injury and oxidative damage. *Plant Physiology*, 109, 1159-1166.
- LEGG, J. P. & FAUQUET, C. M. 2004. Cassava mosaic geminiviruses in Africa. *Plant Molecular Biology*, 56, 585-599.
- LI, Y., DAI, Y., LIU, M., PAN, R., LUO, Y., XIA, Y. & XIA, X. 2009. Scopoletin induces apoptosis of fibroblast-like synoviocytes from adjuvant arthritis rats by a mitochondrial-dependent pathway. *Drug Development Research*, 70, 378-385.
- LIN, H. C., TSAI, S. H., CHEN, C. S., CHANG, Y. C., LEE, C. M., LAI, Z. Y. & LIN, C. M. 2008. Structure-activity relationship of coumarin derivatives on xanthine oxidase-inhibiting and free radical-scavenging activities. *Biochemical Pharmacology*, 75, 1416-25.
- LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative pcr and the  $2^{-\Delta\Delta Ct}$  method. *Methods*, 25, 402-8.
- LOZANO, J. C., TORO, J. C., CASTRO, A. & BELLOTTI, A. C. 1977. *Production of cassava planting material*, Cali, Centro Internacional de Agricultura Tropical (CIAT).
- LOZANO, J. C. A. T., J.C. AND CASTRO, A. AND BELLOTTI, A.C. 1977. *Production of cassava planting material*, Cali, Centro Internacional de Agricultura Tropical (CIAT).
- LULAI, E. C. & SUTTLE, J. C. 2004. The involvement of ethylene in wound-induced suberization of potato tuber (*Solanum tuberosum* L.): a critical assessment. *Postharvest Biology and Technology*, 34, 105-112.
- LULAI, E. C., SUTTLE, J. C. & PEDERSON, S. M. 2008. Regulatory involvement of abscisic acid in potato tuber wound-healing. *Journal of Experimental Botany*, 59, 1175-86.
- MAEDA, H., YOO, H. & DUDAREVA, N. 2011. Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate. *Nature Chemical Biology*, 7, 19-21.
- MANU-ADUENING, J. A., LAMBOLL, R. I., DANKYI, A. A. & GIBSON, R. W. 2005. Cassava diversity in Ghanaian farming systems. *Euphytica*, 144, 331-340.



- MANUELE, M. G., FERRARO, G., BARREIRO ARCOS, M. L., LOPEZ, P., CREMASCHI, G. & ANESINI, C. 2006. Comparative immunomodulatory effect of scopoletin on tumoral and normal lymphocytes. *Life Sciences*, 79, 2043-8.
- MARUTHI, M. N., HILLOCKS, R. J., MTUNDA, K., RAYA, M. D., MUHANNA, M., KIOZIA, H., REKHA, A. R., COLVIN, J. & THRESH, J. M. 2005. Transmission of cassava brown streak virus by *Bemisia tabaci* (Gennadius). *Journal of Phytopathology*, 153, 307-312.
- MATZKE, A. J. & MATZKE, M. A. 1998. Position effects and epigenetic silencing of plant transgenes. *Current Opinion in Plant Biology*, 1, 142-8.
- MATZKE, M. A., MATZKE, A. J., PRUSS, G. J. & VANCE, V. B. 2001. RNA-based silencing strategies in plants. *Current Opinion in Genetics & Development*, 11, 221-7.
- MCWILLIAM, H., LI, W., ULUDAG, M., SQUIZZATO, S., PARK, Y. M., BUSO, N., COWLEY, A. P. & LOPEZ, R. 2013. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Research*, 41, W597-600.
- MEDOUA, G., LAPEMBOME, I., AGBOREGBE, T. & MBOFUNG, C. 2007. Antinutritional factors changes occurring in trifoliate yam (*Dioscorea dumetorum*) tubers after harvest. *Food Chemistry*, 102, 716-720.
- MIFLIN, B. J. & LEA, P. J. 1977. Amino acid metabolism. *Annual Review of Plant Physiology*, 28, 299-329.
- MILLER, R. W., SIROIS, J. C. & MORITA, H. 1975. The reaction of coumarins with horseradish peroxidase. *Plant Physiology*, 55, 35-41.
- MKPONG, O. E., YAN, H., CHISM, G. & SAYRE, R. T. 1990. Purification, characterization, and localization of linamarase in cassava. *Plant Physiology*, 93, 176-81.
- MKUMBIRA, J., CHIWONA-KARLTUN, L., LAGERCRANTZ, U., MAHUNGU, N. M., SAKA, J., MHONE, A., BOKANGA, M., BRIMER, L., GULLBERG, U. & ROSLING, H. 2003. Classification of cassava into 'bitter' and 'cool' in Malawi: From farmers' perception to characterisation by molecular markers. *Euphytica*, 132, 7-22.
- MORANT, A. V., JORGENSEN, K., JORGENSEN, C., PAQUETTE, S. M., SANCHEZ-PEREZ, R., MOLLER, B. L. & BAK, S. 2008.  $\beta$ -Glucosidases as detonators of plant chemical defense. *Phytochemistry*, 69, 1795-813.
- MORENO, I., GRUISSEM, W. & VANDERSCHUREN, H. 2011. Reference genes for reliable potyvirus quantitation in cassava and analysis of cassava brown streak virus load in host varieties. *Journal of Virological Methods*, 177, 49-54.
- MWENJE, E., RIDE, J. P. & PEARCE, R. B. 1998. Distribution of Zimbabwean *Armillaria* groups and their pathogenicity on cassava. *Plant Pathology*, 47, 623-634.
- NAPOLI, C., LEMIEUX, C. & JORGENSEN, R. 1990. Introduction of a chimeric chalcone synthase gene into *Petunia* results in reversible co-suppression of homologous genes in trans. *Plant Cell*, 2, 279-289.
- NASSAR, N. & ORTIZ, R. 2010. Breeding cassava to feed the poor. *Scientific American*, 302, 78-82, 84.
- NEGROTTA, D., JOLLEY, M., BEER, S., WENCK, A. R. & HANSEN, G. 2000. The use of phosphomannose-isomerase as a selectable marker to recover transgenic maize plants (*Zea mays* L.) via *Agrobacterium* transformation. *Plant Cell Reports*, 19, 798-803.
- NGUYEN, T. L., GHEEWALA, S. H. & GARIVAIT, S. 2007. Full chain energy analysis of fuel ethanol from cassava in Thailand. *Environmental Science & Technology*, 41, 4135-42.
- OEY, M., LOHSE, M., KREIKEMEYER, B. & BOCK, R. 2009. Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. *The Plant Journal*, 57, 436-45.
- OHADIKE, D. C. 2009. The influenza pandemic of 1918–19 and the spread of cassava cultivation on the lower Niger: a study in historical linkages. *The Journal of African History*, 22, 379.
- OHAZURIKE, N. C. & ARINZE, A. E. 1996. Changes in phenol oxidase and peroxidase levels in cocoyam tubers of different postharvest ages infected by *Sclerotium rolfsii* sacc. *Food / Nahrung*, 40, 25-27.

- OJALA, T., REMES, S., HAANSUU, P., VUORELA, H., HILTUNEN, R., HAAHTELA, K. & VUORELA, P. 2000. Antimicrobial activity of some coumarin containing herbal plants growing in Finland. *Journal of Ethnopharmacology*, 73, 299-305.
- OLSEN, K. M. & SCHAAL, B. A. 1999. Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences USA*, 96, 5586-5591.
- OLSEN, K. M. & SCHAAL, B. A. 2001. Microsatellite variation in cassava (*Manihot esculenta*, Euphorbiaceae) and its wild relatives: further evidence for a southern Amazonian origin of domestication. *American Journal of Botany*, 88, 131-142.
- OLSON, M. M. & ROSELAND, C. R. 1991. Induction of the coumarins scopoletin and ayapin in sunflower by insect-feeding stress and effects of coumarins on the feeding of sunflower beetle (Coleoptera: Chrysomelidae). *Environmental Entomology*, 20, 1166-1172.
- ONS, O. F. N. S. 2014. *Family food method RNI* [Online]. Available: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/384775/familyfood-method-rni-11dec14.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/384775/familyfood-method-rni-11dec14.pdf) [Accessed 15/11 2015].
- ONWUEME, I. C. 1978. *The tropical tuber crops*, Chichester, John Wiley and Sons Ltd.
- ONWUEME, I. C. 2002. Cassava in Asia and the Pacific. In: HILLOCKS, R. J., THRESH, J. M. & BELLOTTI, A. (eds.) *Cassava : biology, production and utilization*. Wallingford: CABI Pub.
- OSSOWSKI, S., SCHWAB, R. & WEIGEL, D. 2008. Gene silencing in plants using artificial microRNAs and other small RNAs. *The Plant Journal*, 53, 674-90.
- OWITI, J., GROSSMANN, J., GEHRIG, P., DESSIMOZ, C., LALOI, C., HANSEN, M. B., GRUISSEM, W. & VANDERSCHUREN, H. 2011. iTRAQ-based analysis of changes in the cassava root proteome reveals pathways associated with post-harvest physiological deterioration. *The Plant Journal*, 67, 145-56.
- P REZ, F. J. & ORME O-NU EZ, J. 1991. Root exudates of wild oats: Allelopathic effect on spring wheat. *Phytochemistry*, 30, 2199-2202.
- PAGE, M. 2009. *Modulation of root antioxidant status to delay cassava post-harvest physiological deterioration.*, University of Bath.
- PEACH, C. & VELTEN, J. 1991. Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. *Plant Molecular Biology*, 17, 49-60.
- PERONI, N., KAGEYAMA, P. Y. & BEGOSSI, A. 2007. Molecular differentiation, diversity, and folk classification of "sweet" and "bitter" cassava (*Manihot esculenta*) in Caiçara and Caboclo management systems (Brazil). *Genetic Resources and Crop Evolution*, 54, 1333-1349.
- PIPERNO, D. R. & HOLST, I. 1998. The presence of starch grains on prehistoric stone tools from the humid neotropics: Indications of early tuber use and agriculture in Panama. *Journal of Archaeological Science*, 25, 765-776.
- PLUMBLEY, R. A., HUGHES, P. A. & MARRIOTT, J. 1981. Studies on peroxidases and vascular discoloration in cassava root tissue. *Journal of the Science of Food and Agriculture*, 32, 723-731.
- POLISENO, L., SALMENA, L., ZHANG, J., CARVER, B., HAVEMAN, W. J. & PANDOLFI, P. P. 2010. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature*, 465, 1033-8.
- PREUSS, S. & PIKAARD, C. S. 2004. Targeted gene silencing in plants using rna interference. In: ENGELKE, D. R. (ed.) *RNA Interference (RNAi): The Nuts & Bolts of RNAi Technology (Nuts & Bolts Series)*. Dna Press.
- QI, Y., DENLI, A. M. & HANNON, G. J. 2005. Biochemical specialization within *Arabidopsis* RNA silencing pathways. *Molecular Cell*, 19, 421-8.
- RAUHA, J. P., REMES, S., HEINONEN, M., HOPIA, A., KAHKONEN, M., KUJALA, T., PIHLAJA, K., VUORELA, H. & VUORELA, P. 2000. Antimicrobial effects of Finnish plant extracts containing flavonoids and other phenolic compounds. *International Journal of Food Microbiology*, 56, 3-12.
- RAVI, V. & AKED, J. 1996. Review on tropical root and tuber crops. II. Physiological disorders in freshly stored roots and tubers. *Critical Reviews in Food Science and Nutrition*, 36, 711-31.

- RAVI, V., AKED, J. & BALAGOPALAN, C. 1996. Review on tropical root and tuber crops. I. Storage methods and quality changes. *Critical Reviews in Food Science and Nutrition*, 36, 661-709.
- RAY, R. C. & RAVI, V. 2005. Post harvest spoilage of sweetpotato in tropics and control measures. *Critical Reviews in Food Science and Nutrition*, 45, 623-44.
- RE, R., PELLEGRINI, N., PROTEGGENTE, A., PANNALA, A., YANG, M. & RICE-EVANS, C. 1999. Antioxidant activity applying an improved ABTS radical cation decolorization assay. *Free Radical Biology and Medicine*, 26, 1231-7.
- REILLY, K., BERNAL, D., CORTES, D. F., GOMEZ-VASQUEZ, R., TOHME, J. & BEECHING, J. R. 2007. Towards identifying the full set of genes expressed during cassava post-harvest physiological deterioration. *Plant Molecular Biology*, 64, 187-203.
- REILLY, K., GOMEZ-VASQUEZ, R., BUSCHMANN, H., TOHME, J. & BEECHING, J. R. 2003. Oxidative stress responses during cassava post-harvest physiological deterioration. *Plant Molecular Biology*, 53, 669-685.
- REUVENI, M. & COHEN, Y. 1978. Growth retardation and changes in phenolic compounds, with special reference to scopoletin, in mildewed and ethylene-treated tobacco plants. *Physiological Plant Pathology*, 12, 179-189.
- REYES, L. F. & CISNEROS-ZEVALLOS, L. 2003. Wounding stress increases the phenolic content and antioxidant capacity of purple-flesh potatoes (*Solanum tuberosum* L.). *Journal of Agricultural and Food Chemistry*, 51, 5296-300.
- RHODES, J. H. & WOOLTORTON, L. S. C. 1978. The biosynthesis of phenolic compounds in wounded plant storage tissues. . In: KAHL, G. (ed.) *Biochemistry of Wounded Plant Tissues*. Berlin: Walter de Gruyter.
- RICKARD, J. E. 1985. Physiological deterioration of cassava roots. *Journal of the Science of Food and Agriculture*, 36, 167-176.
- RICKARD, J. E. & GAHAN, P. B. 1983. The development of occlusions in cassava (*Manihot esculenta* crantz) root xylem vessels. *Annals of Botany*, 52, 811-821.
- ROCHA-SOSA, M., UWE SONNEWALD, WOLF FROMMER, MARINA STRATMANN, JEFF SCHELL & WILLMITZER, L. 1989. Both developmental and metabolic signals activate the promoter of a class I patatin gene. *The EMBO Journal*, 8, 23-29.
- RODRIGUEZ-CELMA, J., LIN, W. D., FU, G. M., ABADIA, J., LOPEZ-MILLAN, A. F. & SCHMIDT, W. 2013. Mutually exclusive alterations in secondary metabolism are critical for the uptake of insoluble iron compounds by *Arabidopsis* and *Medicago truncatula*. *Plant Physiology*, 162, 1473-85.
- SAHOO, M. R., KOLE, P. C., DASGUPTA, M. & MUKHERJEE, A. 2009. Changes in phenolics, polyphenol oxidase and its isoenzyme patterns in relation to resistance in taro against *Phytophthora colocasiae*. *Journal of Phytopathology*, 157, 145-153.
- SAITOU, N. & NEI, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4, 406-25.
- SAKAI, T., NAKAGAWA-MURATA, Y., URITANI, I. & DATA, E. S. 1994. Occurrence and characteristics of stress metabolites in cassava roots. In: URITANI, I., GARCIA, V. V. & MENDOZA, E. M. T. (eds.) *Postharvest biochemistry of plant food-materials in the tropics*. Tokyo: Japan scientific societies press.
- SAKAI, T. & NAKAGAWA, Y. 1988. Diterpenic stress metabolites from cassava roots. *Phytochemistry*, 27, 3769-3779.
- SAKAI, T., NAKAGAWA, Y., URITANI, I. & DATA, E. S. 1986. Occurrence of various kinds of metabolites in physiologically and microbially damaged cassava (*Manihot esculenta* Crantz) roots. *Agricultural and Biological Chemistry*, 50, 2905-2907.
- SAMBATTI, J. B. M., MARTINS, P. S. & ANDO, A. 2001. Folk taxonomy and evolutionary dynamics of cassava: A case study in Ubatuba, Brazil. *Economic Botany*, 55, 93-105.
- SAMBROOK, J., FRITSCH, E. F. & MAMIATIS, T. 1989. Molecular cloning: A laboratory manual. 2nd ed. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- SANTANA, M. A., VASQUEZ, V., MATEHUS, J. & ALDAO, R. R. 2002. Linamarase expression in cassava cultivars with roots of low- and high-cyanide content. *Plant Physiology*, 129, 1686-94.

- SARDARI, S., NISHIBE, S. & DANESHTALAB, M. 2000. Coumarins, the bioactive structures with antifungal property. *Studies in Natural Products Chemistry*, 23, 335-393.
- SAYRE, R., BEECHING, J. R., CAHOON, E. B., EGESI, C., FAUQUET, C., FELLMAN, J., FREGENE, M., GRUISSEM, W., MALLOWA, S., MANARY, M., MAZIYA-DIXON, B., MBANASO, A., SCHACHTMAN, D. P., SIRITUNGA, D., TAYLOR, N., VANDERSCHUREN, H. & ZHANG, P. 2011. The BioCassava plus program: biofortification of cassava for sub-Saharan Africa. *Annual Review of Plant Biology*, 62, 251-72.
- SCHADEL, W. E. & WALTER, W. M. J. 1981. Localization of phenols and polyphenol oxidase in 'Jewel' sweet potatoes (*Ipomoea batatas* 'Jewel'). *Canadian Journal of Botany*, 59, 1961-1967.
- SCHILMILLER, A. L. & HOWE, G. A. 2005. Systemic signaling in the wound response. *Current Opinion in Plant Biology*, 8, 369-77.
- SCHMID, M., DAVISON, T. S., HENZ, S. R., PAPE, U. J., DEMAR, M., VINGRON, M., SCHOLKOPF, B., WEIGEL, D. & LOHMANN, J. U. 2005. A gene expression map of *Arabidopsis thaliana* development. *Nature Genetics*, 37, 501-6.
- SCHMID, N. B., GIEHL, R. F., DOLL, S., MOCK, H. P., STREHMEL, N., SCHEEL, D., KONG, X., HIDER, R. C. & VON WIREN, N. 2014. Feruloyl-CoA 6'-Hydroxylase1-dependent coumarins mediate iron acquisition from alkaline substrates in *Arabidopsis*. *Plant Physiology*, 164, 160-72.
- SCHOCH, G., GOEPFERT, S., MORANT, M., HEHN, A., MEYER, D., ULLMANN, P. & WERCK-REICHHART, D. 2001. CYP98A3 from *Arabidopsis thaliana* is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *The Journal of Biological Chemistry*, 276, 36566-74.
- SCHULZE-SIEBERT, D., HEINEKE, D., SCHARF, H. & SCHULTZ, G. 1984. Pyruvate-derived amino acids in spinach chloroplasts : synthesis and regulation during photosynthetic carbon metabolism. *Plant Physiology*, 76, 465-471.
- SCHWAB, R., OSSOWSKI, S., RIESTER, M., WARTHMAN, N. & WEIGEL, D. 2006. Highly specific gene silencing by artificial microRNAs in *Arabidopsis*. *Plant Cell*, 18, 1121-33.
- SCHWAB, R., PALATNIK, J. F., RIESTER, M., SCHOMMER, C., SCHMID, M. & WEIGEL, D. 2005. Specific effects of microRNAs on the plant transcriptome. *Developmental Cell*, 8, 517-527.
- SEAL, H. L. 1967. Studies in the history of probability and statistics. XV: The historical development of the Gauss linear model. *Biometrika*, 54, 1.
- SHARMA, P., JHA, A. B., DUBEY, R. S. & PESSARAKLI, M. 2012. Reactive oxygen species, oxidative damage, and antioxidative defense mechanism in plants under stressful conditions. *Journal of Botany*, 2012, 1-26.
- SHAW, C. Y., CHEN, C. H., HSU, C. C., CHEN, C. C. & TSAI, Y. C. 2003. Antioxidant properties of scopoletin isolated from *Sinomonium acutum*. *Phytotherapy Research*, 17, 823-5.
- SHUKLA, Y. N., SRIVASTAVA, A., KUMAR, S. & KUMAR, S. 1999. Phytotoxic and antimicrobial constituents of *Argyrea speciosa* and *Oenothera biennis*. *Journal of Ethnopharmacology*, 67, 241-245.
- SIMON, H. U., HAJ-YEHIA, A. & LEVI-SCHAFFER, F. 2000. Role of reactive oxygen species (ROS) in apoptosis induction. *Apoptosis*, 5, 415-418.
- SIRITUNGA, D. & SAYRE, R. T. 2003. Generation of cyanogen-free transgenic cassava. *Planta*, 217, 367-73.
- SIWINSKA, J., KADZINSKI, L., BANASIUK, R., GWIZDEK-WISNIEWSKA, A., OLRZY, A., BANECKI, B., LOJKOWSKA, E. & IHNATOWICZ, A. 2014. Identification of QTL affecting scopolin and scopoletin biosynthesis in *Arabidopsis thaliana*. *BMC Plant Biology*, 14, 280.
- SMITH, N. A., SINGH, S. P., WANG, M., STOUTJESDIJK, P. A., GREEN, A. G. & WATERHOUSE, P. M. 2000. Total silencing by intronsplined hairpin RNAs. *Nature*, 407, 319-320.
- SOLOMOS, T. & LATIES, G. G. 1976. Induction by ethylene of cyanide-resistant respiration. *Biochemical and Biophysical Research Communications*, 70, 663-671.
- STAHMANN, M. A., CLARE, B. G. & WOODBURY, W. 1966. Increased disease resistance and enzyme activity induced by ethylene and ethylene production of black rot infected sweet potato tissue. *Plant Physiology*, 41, 1505-1512.

- STUPAK, M., VANDERSCHUREN, H., GRUISSEM, W. & ZHANG, P. 2006. Biotechnological approaches to cassava protein improvement. *Trends in Food Science & Technology*, 17, 634-641.
- SUDHA, G. & RAVISHANKAR, G. A. 2002. Involvement and interaction of various signaling compounds on the plant metabolic events during defense response, resistance to stress factors, formation of secondary metabolites and their molecular aspects. *Plant Cell Tissue and Organ Culture*, 71, 181-212.
- SUN, H., WANG, L., ZHANG, B., MA, J., HETTENHAUSEN, C., CAO, G., SUN, G., WU, J. & WU, J. 2014. Scopoletin is a phytoalexin against *Alternaria alternata* in wild tobacco dependent on jasmonate signalling. *Journal of Experimental Botany*, 65, 4305-15.
- TANAKA, Y., DATA, E. S., LAPE, V. G., VILLEGAS, C. D., GORGONIO, M., HIROSE, S. & URITANI, I. 1984. Effect of pruning treatment on physiological deterioration in cassava roots. *Agricultural and Biological Chemistry*, 48, 739-743.
- THUONG, P. T., HUNG, T. M., NGOC, T. M., HA DO, T., MIN, B. S., KWACK, S. J., KANG, T. S., CHOI, J. S. & BAE, K. 2010. Antioxidant activities of coumarins from Korean medicinal plants and their structure-activity relationships. *Phytotherapy Research*, 24, 101-6.
- TIAN, X., XIE, J., ZHAO, Y., LU, H., LIU, S., QU, L., LI, J., GAI, Y. & JIANG, X. 2013. Sense-, antisense- and RNAi-4CL1 regulate soluble phenolic acids, cell wall components and growth in transgenic *Populus tomentosa* Carr. *Plant Physiology and Biochemistry*, 65, 111-9.
- TIKU, M. L. 1971. Power function of the F-test under non-normal situations. *Journal of the American Statistical Association*, 66, 913-916.
- TOHGE, T., WATANABE, M., HOEFGEN, R. & FERNIE, A. R. 2013. The evolution of phenylpropanoid metabolism in the green lineage. *Critical Reviews in Biochemistry and Molecular Biology*, 48, 123-52.
- TOM S-BARBER N, F. A. & ESP N, J. C. 2001. Phenolic compounds and related enzymes as determinants of quality in fruits and vegetables. *Journal of the Science of Food and Agriculture*, 81, 853-876.
- TOUFIGHI, K., BRADY, S. M., AUSTIN, R., LY, E. & PROVART, N. J. 2005. The Botany Array Resource: e-Northern, Expression Angling, and promoter analyses. *The Plant Journal*, 43, 153-63.
- TRIPATHI, A. K., BHAKUNI, R. S., UPADHYAY, S. & GAUR, R. 2011. Insect feeding deterrent and growth inhibitory activities of scopoletin isolated from *Artemisia annua* against *Spilarctia obliqua* (Lepidoptera: Noctuidae). *Insect Science*, 18, 189-194.
- TUSCHL, T., ZAMORE, P. D., LEHMANN, R., BARTEL, D. P. & SHARP, P. A. 1999. Targeted mRNA degradation by double-stranded RNA in vitro. *Genes & Development*, 13, 3191-7.
- UARROTA, V. G., MORESCO, R., COELHO, B., NUNES EDA, C., PERUCH, L. A., NEUBERT EDE, O., ROCHA, M. & MARASCHIN, M. 2014. Metabolomics combined with chemometric tools (PCA, HCA, PLS-DA and SVM) for screening cassava (*Manihot esculenta* Crantz) roots during postharvest physiological deterioration. *Food Chemistry*, 161, 67-78.
- UGENT, D., POZORSKI, S. & POZORSKI, T. 1986. Archaeological manioc (*Manihot*) from coastal peru. *Economic Botany*, 40, 78-102.
- VAN BREUSEGEM, F. & DAT, J. F. 2006. Reactive oxygen species in plant cell death. *Plant Physiology*, 141, 384-90.
- VAN OIRSCHOT, Q. E. A., O'BRIEN, G. M., DUFOUR, D., EL-SHARKAWY, M. A. & MESA, E. 2000. The effect of pre-harvest pruning of cassava upon root deterioration and quality characteristics. *Journal of the Science of Food and Agriculture*, 80, 1866-1873.
- VANDERSCHUREN, H., AKBERGENOV, R., POOGGIN, M. M., HOHN, T., GRUISSEM, W. & ZHANG, P. 2007a. Transgenic cassava resistance to African cassava mosaic virus is enhanced by viral DNA-A bidirectional promoter-derived siRNAs. *Plant Molecular Biology*, 64, 549-57.
- VANDERSCHUREN, H., NYABOGA, E., POON, J. S., BAERENFALLER, K., GROSSMANN, J., HIRSCH-HOFFMANN, M., KIRCHGESSNER, N., NANNI, P. & GRUISSEM, W. 2014. Large-scale proteomics of the cassava storage root and identification of a target gene to reduce postharvest deterioration. *Plant Cell*, 26, 1913-1924.

- VANDERSCHUREN, H., STUPAK, M., FUTTERER, J., GRUISSEM, W. & ZHANG, P. 2007b. Engineering resistance to geminiviruses -- review and perspectives. *Plant Biotechnology Journal*, 5, 207-20.
- VANIN, E. F. 1985. Processed pseudogenes: characteristics and evolution. *Annual Review of Genetics*, 19, 253-72.
- VINCZE, T., POSFAI, J. & ROBERTS, R. J. 2003. NEBcutter: A program to cleave DNA with restriction enzymes. *Nucleic Acids Research*, 31, 3688-91.
- VOGEL, C. & MARCOTTE, E. M. 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics*, 13, 227-32.
- VOGT, T. 2010. Phenylpropanoid biosynthesis. *Molecular Plant*, 3, 2-20.
- WANG, W., FENG, B., XIAO, J., XIA, Z., ZHOU, X., LI, P., ZHANG, W., WANG, Y., MOLLER, B. L., ZHANG, P., LUO, M. C., XIAO, G., LIU, J., YANG, J., CHEN, S., RABINOWICZ, P. D., CHEN, X., ZHANG, H. B., CEBALLOS, H., LOU, Q., ZOU, M., CARVALHO, L. J., ZENG, C., XIA, J., SUN, S., FU, Y., WANG, H., LU, C., RUAN, M., ZHOU, S., WU, Z., LIU, H., KANNANGARA, R. M., JORGENSEN, K., NEALE, R. L., BONDE, M., HEINZ, N., ZHU, W., WANG, S., ZHANG, Y., PAN, K., WEN, M., MA, P. A., LI, Z., HU, M., LIAO, W., HU, W., ZHANG, S., PEI, J., GUO, A., GUO, J., ZHANG, J., ZHANG, Z., YE, J., OU, W., MA, Y., LIU, X., TALLON, L. J., GALENS, K., OTT, S., HUANG, J., XUE, J., AN, F., YAO, Q., LU, X., FREGENE, M., LOPEZ-LAVALLE, L. A., WU, J., YOU, F. M., CHEN, M., HU, S., WU, G., ZHONG, S., LING, P., CHEN, Y., WANG, Q., LIU, G., LIU, B., LI, K. & PENG, M. 2014. Cassava genome from a wild ancestor to cultivated varieties. *Nature Communications*, 5, 5110.
- WATERHOUSE, P. M., WANG, M.-B. & LOUGH, T. 2001. Gene silencing as an adaptive defence against viruses. *Nature*, 411, 834-842.
- WESLEY, S. V., HELLIWELL, C. A., SMITH, N. A., WANG, M. B., ROUSE, D. T., LIU, Q., GOODING, P. S., SINGH, S. P., ABBOTT, D., STOUTJESDIJK, P. A., ROBINSON, S. P., GLEAVE, A. P., GREEN, A. G. & WATERHOUSE, P. M. 2001. Construct design for efficient, effective and high-throughput gene silencing in plants. *The Plant Journal*, 27, 581-90.
- WHEATLEY, C. 1982. *Studies on cassava (Manihot esculenta Crantz) root physiological deterioration*. PhD, University of London.
- WHEATLEY, C. C. & SCHWABE, W. W. 1985. Scopoletin involvement in post-harvest physiological deterioration of cassava root (*Manihot esculenta* Crantz). *Journal of Experimental Botany*, 36, 783-791.
- WILKINSON, J. E., TWELL, D. & LINDSEY, K. 1997. Activities of CaMV 35S and nos promoters in pollen: implications for field release of transgenic plants. *Journal of Experimental Botany*, 48, 265-275.
- WILSON, W. M. & DUFOUR, D. L. 2002. Why "bitter" cassava? Productivity of "bitter" and "sweet" cassava in a Tukanoan Indian settlement in the northwest Amazon. *Economic Botany*, 56, 49-57.
- WINKEL-SHIRLEY, B. 2002a. Biosynthesis of flavonoids and effects of stress. *Current Opinion in Plant Biology*, 5, 218-23.
- WINKEL-SHIRLEY, B. 2002b. Biosynthesis of flavonoids and effects of stress. *Current Opinion in Plant Biology*, 5, 218-23.
- WINTER, D., VINEGAR, B., NAHAL, H., AMMAR, R., WILSON, G. V. & PROVART, N. J. 2007. An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS One*, 2, e718.
- WOJTASZEK, P. 1997. Oxidative burst: an early plant response to pathogen infection. *Biochemical Journal*, 322 ( Pt 3), 681-92.
- XU, J., DUAN, X., YANG, J., BEECHING, J. R. & ZHANG, P. 2013. Enhanced reactive oxygen species scavenging by overproduction of superoxide dismutase and catalase delays postharvest physiological deterioration of cassava storage roots. *Plant Physiology*, 161, 1517-28.
- XU, P., ZHANG, Y., KANG, L., ROOSSINCK, M. J. & MYSORE, K. S. 2006. Computational estimation and experimental verification of off-target silencing during posttranscriptional gene silencing in plants. *Plant Physiology*, 142, 429-40.

- YANG, J., AN, D. & ZHANG, P. 2011. Expression profiling of cassava storage roots reveals an active process of glycolysis/gluconeogenesis. *Journal of Integrative Plant Biology*, 53, 193–211.
- YE, J., COULOURIS, G., ZARETSKAYA, I., CUTCUTACHE, I., ROZEN, S. & MADDEN, T. L. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, 13, 134.
- YEOH, H. H. 1996. Protein contents, amino acid compositions and nitrogen-to-protein conversion factor for cassava roots. *Journal of the Science of Food and Agriculture*, 70, 51-54.
- YU, J., GUO, Y., GUO, A. & HE, L. 2008. 木薯块根总 RNA 提取方法的比较和改进. *Journal of Anhui Agricultural Science*, 36, 4872-4873.
- ZHANG, C., HAN, W. J., JING, X. D., PU, G. Q. & WANG, C. T. 2003a. Life cycle economic analysis of fuel ethanol derived from cassava in southwest China. *Renewable & Sustainable Energy Reviews*, 7, 353-366.
- ZHANG, P., BOHL-ZENGER, S., PUONTI-KAERLAS, J., POTRYKUS, I. & GRUISSEM, W. 2003b. Two cassava promoters related to vascular expression and storage root formation. *Planta*, 218, 192-203.
- ZHANG, P., VANDERSCHUREN, H., FUTTERER, J. & GRUISSEM, W. 2005. Resistance to cassava mosaic disease in transgenic cassava expressing antisense RNAs targeting virus replication genes. *Plant Biotechnology Journal*, 3, 385-97.
- ZHANG, X., HENRIQUES, R., LIN, S. S., NIU, Q. W. & CHUA, N. H. 2006. *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nature Protocols*, 1, 641-6.
- ZIDENGA, T., LEYVA-GUERRERO, E., MOON, H., SIRITUNGA, D. & SAYRE, R. 2012. Extending cassava root shelf life via reduction of reactive oxygen species production. *Plant Physiology*, 159, 1396-407.
- ZISKA, L. H., RUNION, G. B., TOMECEK, M., PRIOR, S. A., TORBET, H. A. & SICHER, R. 2009. An evaluation of cassava, sweet potato and field corn as potential carbohydrate sources for bioethanol production in Alabama and Maryland. *Biomass & Bioenergy*, 33, 1503-1508.

## 8 APPENDIX

### APPENDIX I PRIMERS AND THEIR DETAILS

Name	Sequence (5'to3')	Tm (°C)
>salk129938_LP	GGTCGGGATTCTAATCTCAGC	60.05
>salk129938_RP	AGAAGATGGTGAGGAGGCTTC	59.83
>salk5232_LP	CATCCCAAAGTCGATGATCAC	60.33
>salk5232_RP	TTGTGATGCTGCTGAGAAATG	60
>Ath-13610-F	ATGGCTCCAACACTCTTGACAACC	63.33
>Ath-13610-F2	AGCCTCTTCTTTGTCTCTGAAGC	60.56
>Ath-13610-f3-hi gh	TGATGACGTCAAGGCTGCGA	62.15
>Ath-13610-R	TTCTTTCCATCGTGTGCCTTCCTG	63.37
>SALK-LBb1	GCGTGGACCGCTTGCTGCAACT	69.3
>SAIL-LB3	TAGCATCTGAATTCATAACCAATCTCG ATACAC	62.2
>010292-g-outer- fwd	GAACGGAATCTTTGTCGTCATTGAG	60.67
>010292-g-outer- rvs	GGCTTTAGTCACAGGCCGAA	60.32
>010376-c-fwd	AAAAGCAGGCTCCACCCTCAAAAATCAA CGTTTC	low: 58.9, high: 67.8
>010376-c-rvs	GAAAGCTGGGTGGAGAAAATTTATCACA CTTCAGCA	58.1, 66.8
>010292-g-inner- fwd	AAAAGCAGGCTTGTGCTCTCATTCTCC CTACC	59.5, 68.3
>010292-g-inner- rvs	GAAAGCTGGGTCTTGGACGATACATTTG ACACATC	58.2, 66.9
>010291-g -fwd	AAAAGCAGGCTCCTATCTCAAACAATAT TACCTGTGAGG	59.6, 66.9
>010291-g -rvs	GAAAGCTGGGTGGCATGTCCACATTAT CAAGC	59.3, 68.0
>Attb-FWD	GGGGACAAGTTTGTACAAAAAGCAGG	65.4



	CT	
>Attb-RVS	GGGGACCACTTTGTACAAGAAAGCTGG GT	67.7
>Gus-full-FWD	AAAGCAAGCTGACCGACACT	60.18
>Gus-full-RVS	GCTGGTCACCAATTCACACG	59.76
>GW-full-fwd	TGACGTAAGGGATGACGCAC	60.11
>GW-full-rvs	TAATCATCGCAAGACCGGCA	59.82
>35s-rfa-Fwd	GACGTAAGGGATGACGCACA	60.11
>35s-rfa-Rvs	AGCTCCTGAAAATCTCGACGG	60.13
>PAT-gus-Rvs	ACCGACACTTTCACGGTCTC	59.97
>10376-1exF-fwd	CCAACACTTGCAGAATCAGCC	60.07
>10376-1exF-rvs	ATTAGCCTCGTCGTCGGAGA	60.46
>10376k1-1exF-fwd	AAAAGCAGGCTCCAACACTTGCAGAATC AGCC	low: 68.4 high: 60.07
>10376k1-1exF-rvs	CTGTGGTACCATTAGCCTCGTCGTCGG AGA	low: 68.2 high: 60.46
>10376k2-1exF-fwd	GGGGACAAGTTTGTACAAAAAGCAGG CTCCAACAC	low: 67.9 high: 56.4
>10376k3-1-2exF-fwd	TGGGGAATTCGGGGACAAGTTTGTACA AAAAAGC	low: 66.6 high: 58.2
>10376k1-1exR-fwd	GAAAGCTGGGTCCAACACTTGCAGAAT CAGCC	low: 68.4 high: 60.07
>10376k1-1exR-rvs	TACCCAAGCTTATTAGCCTCGTCGTCGG AGA	low: 67.4 high: 60.46
>10376k2-1exR-fwd	GGGGACCACTTTGTACAAGAAAGCTGG GTCCAA	low: 68.7 high: 51.4
>10376k3-1exR-fwd	TCAGGATCCGGGGACCACTTTGTACAA GAA	low: 66.8 high: 57.8
>pdk-int-fwd	TGCTAATATAACAAAGCGCAAGAT	57.38
>fullseq-pKann-1st-fwd	CCAACCACGTCTTCAAAGCA	58.98
>fullseq-pKann-1st-rvs	TCATCTTACATGTTTCGATCAAATTC	56.04

>fullseq-pKann-2 nd-fwd	TGCTAATATAACAAAGCGCAAGA	56.94
>fullseq-pKann-2 nd-rvs	AACCGGAGTTCCCTTCAGTG	59.6
>csv-ubq10-qpcr -fwd1	CACCGGATCAGCAAAGGCTTA	60
>csv-ubq10-qpcr -rvs1	CAGACACACAGATCAAAGCAGC	60
>3genes-qpcr-fw d1	TGGGTTCATGTTCTCCTCCCAT	60
>3genes-qpcr-rv s1	CTCCTATATCGACCATTGCTGAGTA	60
>hyg-fwd	CCACTATCGGCGAGTACTTCTACACAGC	64.6
>hyg-rvs	GCCTGAACTCACCGCGACGTCTGTC	68.2
>M13-fwd	GTAAAACGACGGCCAGT	55.6
>M13-rvs	AACAGCTATGACCATG	48.4
>10291-f1	GGTACTCGATGGAGTCAAGGAC	59.9
>10291-r1	CGCCGACAGTAAGATTAGGATTG	59.26
>10292-f1	GAGCGTAGAGCATTGTGTGATTA	58.89
>10292-r2	CTTCCTGAAGAAATGCTTGACG	57.92
>10376-f1	TGGGCATCAAAAGCCTTCCTATA	59.8
>10376-f2	GCA ACT CAT GGT TTC TTT GGA C	57.7
>10376-r1	CTGAGGAGGCCACAATGCATT	60.96
>10376-r2	CCACAATGCATTAGCCTCGTCG	62.18
>10381-f1	CAAAGAGCTTTCTTCCACCAACAATA	60.24
>10381-f2	CCTTCCCTCTCAGTATATTCAACCT	59.4
>10381-r1	GTTGATCACAAGGGATCCTTCAATT	59.58
>27567-f1	GCTCCAGCAATGGCAGTG	59.12
>27567-f2	ATTCAACCCCAAGAGGCACTT	59.85
>27567-r1	ACGTTTCAGTTTCTCCATGAGTGT	60.18
>27567-r2	ACGTTTCAGTTTCTCCATGAGTGT	60.18
>27567-r3	AGCTGGTTCCAAATCTGATGT	57.55
>30526-f1	GCATCAAAAGCCTTCCTCGC	60.18
>30526-r1	CACAGCAGCTGGAACCCC	59.16

>30526-r2	CCAAGCATTCATCCTTGCACTC	60.16
>33240-f1	CTTTAGTCCTGATGCAGAGAAG	56.27
>33240-r1	AACAGTCTTCTTGCCATCG	55.54

**Table 8.1 Primers involved in this research project.** Att sequences are in red font; restriction site sequences are in Underlined font.

Number	Pair	Template	Amplicon length (bp)	Working temperature (°C)
1	>M13-fwd	M13 targets on vectors	n/a	55
	>M13-rvs			
2	>hyg-fwd	hptII	978	61
	>hyg-rvs	hygromycin resistant gene		
3	>Ath-13610-F	<i>Arabidopsis</i> gene at3g13610, full length	1144	60
	>Ath-13610-R			
4	>Ath-13610-F2	<i>Arabidopsis</i> gene at3g13610, 3' end	605	57
	>Ath-13610-R			
5	>salk129938_LP	<i>Arabidopsis</i> T-DNA line Salk_129938 T-DNA wild-type region	573-873	57
	>salk129938_RP			
6	>salk5232_LP	<i>Arabidopsis</i> T-DNA line Salk_005232 T-DNA wild-type	570-870	57
	>salk5232_RP			

		region		
<b>7</b>	>salk129938_RP	<i>Arabidopsis</i>	n/a	57
	>SALK-LBb1	T-DNA line Salk_129938 T-DNA insertion		
<b>8</b>	>Ath-13610-R	<i>Arabidopsis</i>	n/a	59
	>SAIL-LB3	T-DNA line Sail_1252_A10 T-DNA insertion		
<b>9</b>	>salk5232_RP	<i>Arabidopsis</i>	n/a	57
	>SALK-LBb1	T-DNA line Salk_005232 T-DNA insertion		
<b>10</b>	>Ath-13610-F2	<i>Arabidopsis</i>	n/a	59
	>SALK-LBb1	T-DNA line Salk_129938 T-DNA insertion in cDNA, alternative		
<b>11</b>	>csv-ubq10-qpcr-fw d1	cassava ubiquitin 10 gene	160	60
	>csv-ubq10-qpcr-rv s1			
<b>12</b>	>3genes-qpcr-fwd1	cassava F6'H1 candidates	86	60
	>3genes-qpcr-rvs1			
<b>13</b>	>010376-c-fwd	cassava gene 10376 + half AttB	1135	low: 55 high: 65
	>010376-c-rvs			
<b>14</b>	>010292-g-outer-fw d	cassava gene 10292 and	2211	57

	>010292-g-outer-rvs	flanking sequence		
<b>15</b>	>010292-g-inner-fw d	cassava gene10292 + half AttB	1277	low: 55 high: 64
	>010292-g-inner-rvs			
<b>16</b>	>010291-g -fwd	cassava gene 10291 + half AttB	1380	low: 56 high: 64
	>010291-g -rvs			
<b>17</b>	>Attb-FWD	AttB sites	n/a	62
	>Attb-RVS			
<b>18</b>	>10376-1exF-fwd	first exon of cassava gene 10376	471	57
	>10376-1exF-rvs			
<b>19</b>	>10376k1-1exF-fwd	product of 18	492	low: 57, high: 65
	>10376k1-1exF-rvs			
<b>20</b>	>10376k2-1exF-fwd	product of 19		low: 57, high: 65
	>10376k1-1exF-rvs			
<b>21</b>	>10376k3-1-2exF-fwd	product of 20		low: 57, high: 65
	>10376k1-1exF-rvs			
<b>22</b>	>10376k1-1exR-fw d	product of 18		low: 57, high: 64
	>10376k1-1exR-rvs			
<b>23</b>	>10376k2-1exR-fw d	product of 22		low: 55, high: 64
	>10376k1-1exR-rvs			
<b>24</b>	>10376k3-1exR-fw d	product of 23		low: 55, high: 64
	>10376k1-1exR-rvs			
<b>25</b>	35s-rfa-Fwd	From CaMV 35S promoter to the initial part of	323	57
	>35s-rfa-Rvs			

		Gateway cassette		
<b>33</b>	>33240-f1	new primers for 33240	670	Variable
	>33240-r1			
<b>34</b>	>10291-f1	new primers for 10291	398	Variable
	>10291-r1			
<b>35</b>	>10292-f1	new primers for 10292	178	Variable
	>10292-r2			
<b>35</b>	>10376-f1	new primers for 10376	401	Variable
	>10376-r1			
<b>36</b>	>10381-f1	new primers for 10381	469	Variable
	>10381-r1			
<b>37</b>	>30526-f1	new primers for 30526	244	Variable
	>30526-r1			
<b>38</b>	>27567-f1	new primers for 27567	578	Variable
	>27567-r1			

**Table 8.2 Primer pairs used in this research project, and their work condition.**

## **APPENDIX II      THE RECIPE OF SOLUTIONS USED IN THIS PROJECT**

### **1, LB medium**

In one litre of distilled water, dissolve

10g Tryptone peptone (Difco)

5g Yeast extract (Becton Dickinson)

10g NaCl

The pH is adjusted to 7.0, and the medium is autoclaved at 121 °C for 15 mins.

If solid agar medium is needed, add 15g agar before autoclave.

### **2, 2YT medium**

In one litre of distilled water, dissolve

16g Tryptone peptone (Difco)

10g Yeast extract (Becton Dickinson)

5g NaCl

The pH is adjusted to 7.0, and the medium is autoclaved at 121 °C for 15 mins.

If solid agar medium is needed, add 15g agar before autoclave.

### **3, SOC medium**

In one litre of distilled water, dissolve

20g Tryptone peptone (Difco)

5g Yeast extract (Becton Dickinson)

0.58g NaCl

0.19g KCl

0.95g MgCl<sub>2</sub>

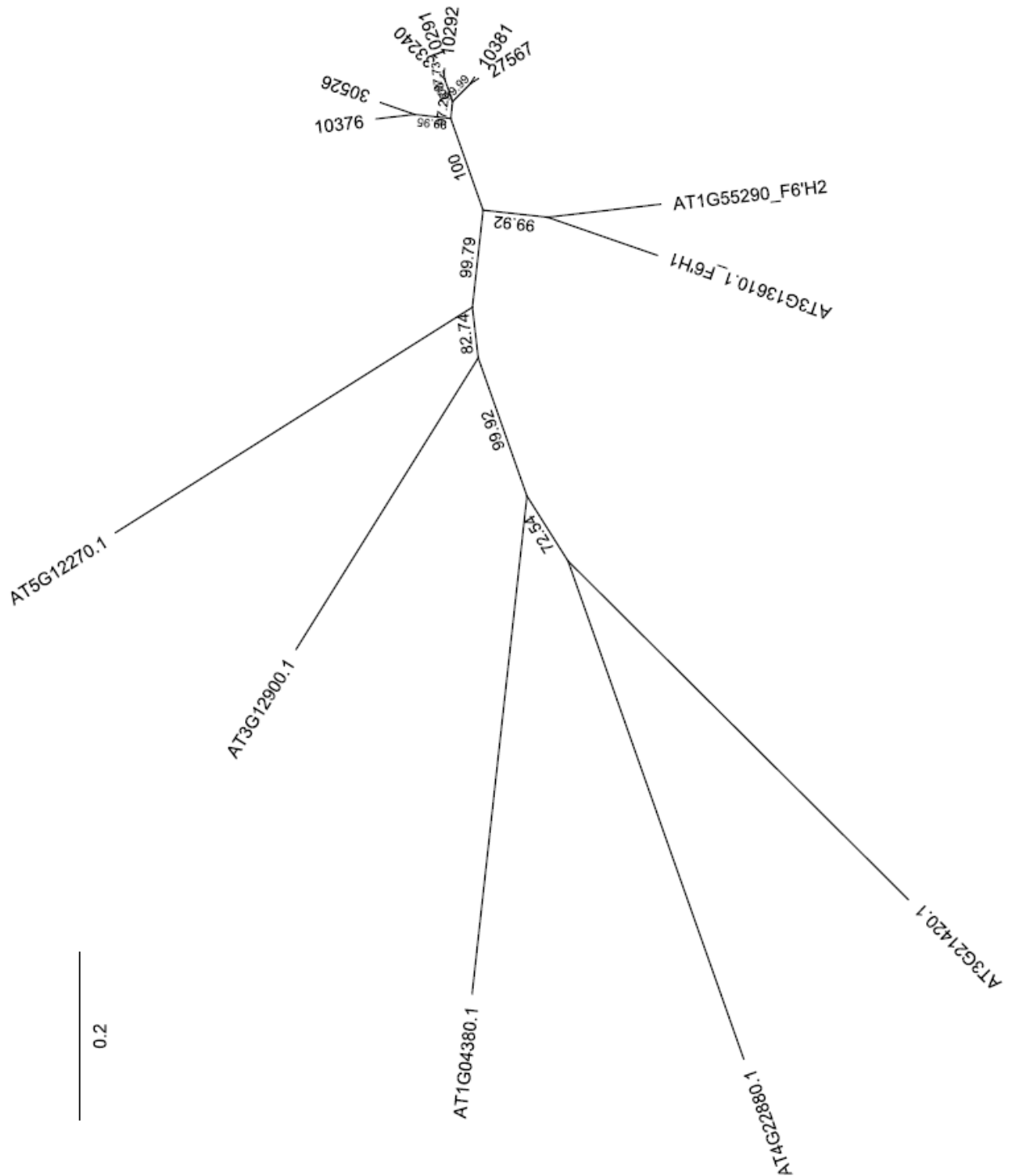
1.2g MgSO<sub>4</sub>

3.6g Glucose

The pH is adjusted to 7, and the medium is autoclaved at 121 °C for 15 mins.



**Fig. 8.1 Enlarged phylogenetic tree of the amino acid sequences of *Arabidopsis* and cassava F6'H1 gene family.**





```

33240  MAPMTAIS—SADSFDLTDFVINKGNVGKGLSDLGKISLPQYIQPQREALNIIPQKSID
10292  MAPMTAVS—SADSFDLTDFVINKGNVGKGLSDLGKISLPQYIQPQREALNIIPQKSID
10291  MAPMTAVS—SADPFDLTDFVINKGNVGKGLSDLGKISLPQYIQPQREALNIIPQKSID
10381  MAPAMALS—STDSFDLTDFVINKGNVGKGLSDLGKISLPQYIQPQREALNIIPQKSID
27567  MAPAMAVS—STDSFDLTDFVINKGNVGKGLSDLGKISLPQYIQPQREALNIIPQKSID
30526  MAPTLAESTTASADFIADFIQNGKGNVGKGLSELGKISLPQYIQPQREALNIIPQKSID
10376  MAPTLAESTA—DSFDLTDFVINKGNVGKGLSDLGKISLPQYIQPQREALNIIPQKSID
***** * : * :***** ***** ***** ***** ***** ***** *****
33240  SKELSTNSVRFGTSFSPDAEKALEWKIDVLSLFYVSEEEAFALWPNACRDEVLEVYMKRSQ
10292  SKELSTNSVRFGTSFSPDAEKALEWKIDVLSLFYVSEEEAFALWPNACRDEVLEVYMKRSQ
10291  SKELSTNSVRFGTSFSPDAEKALEWKIDVLSLFYVSEEEAFALWPNACRDEVLEVYMKRSQ
10381  SKELSTNNIRFGTSFSPDAEKALEWKIDVLSLFYVSEEEASVLWPSACRDEVLEVYMKRSQ
27567  SKELSTNNIRFGTSFSPDAEKALEWKIDVLSLFYVSEEEASVLWPSACRDEVLEVYMKRSQ
27562  SKELSPNSVRFGTSFSPDAEKALEWKIDVLSLFYVSEEEASLWPPCKDECLEVYMKGSE
10376  SKELSPNSVRFGTSFSPDAEKALEWKIDVLSLFYVSDDEAALWPPCKDECLEVYMKGSE
***** * :***** ***** ***** ***** ***** ***** *****
33240  FLIQLDEIGGLVVRVNEKGGEEDGWVHPPIEGSLVINVGDAQLISNGRYRSVEHCIVAS
10292  FLIQLDEIGGLVVRVNEKGGEEDGWVHPPIEGSLVINVGDAQLISNGRYRSVEHCIVTS
10291  FLIQLDEIGGLVVRVNEKGGEEDGWVHPPIEGSLVINVGDAQLISNGRYRSVEHCIVAS
10381  FLIQLDEIGGLVVRVNEKGGEEDGWVHPPIEGSLVINVGDAQLISNGRYRSVEHCIVAS
27567  FLIQLDEIGGLVVRVNEKGGEEDGWVHPPIEGSLVINVGDAQLISNGRYRSVEHCIVAS
27562  FLIQLDEIGGLVVRVNEKGGEEDGWVHPPIEGSLVINVGDAQLISNGRYRSVEHCIVAS
10376  FLIQLDEIGGLVVRVNEKGGEEDGWVHPPIEGSLVINVGDAQLISNGRYRSVEHCIVAS
***** ***** : ***** ***** ***** ***** ***** *****
33240  AEI*—
10292  AEISS*
10291  AEISS*
10381  AEI*—
27567  AEISS*
30526  AEV*—
10376  AEV*—
**

```

**Fig. 8.4 Alignment of the coding DNA sequences of cassava F6'H1 candidate genes.**

```

10376 ATGGCTCCAACTTGCAGAAATCAGCCACTGACTCCCTTTGATCTCACTGATTTTGTAAAT AACCAGGCAATGGCGTTAAGGECCTTTCTGATTTGGGCATCAAAAGCCTTCCTATACAA
27567 ATGGCTCCAGCAATGGCAGTGTATCCACAGATTCCTTTGATCTCACTGATTTTGTCAAT AACAAAGGCAATGGAGTAAAGGTCCTTTCTGATTTAGGCATCAAAAGCCTTCCTCTCAA
10292 ATGGCTCCAACTTGCAGTATCATCAGCTGATCCCTTTGATCTCACTGATTTTGTCAAT AACAAAGGCAATGGAGTAAAGGTCCTTTCTGATTTGGGCATCAAAAGCCTTCCTCTCAA
10291 ATGGCTCCAACTTGCAGTATCATCAGCTGATCCCTTTGATCTCACTGATTTTGTCAAT AACCAGGCAATGGCGTTAAGGECCTTTCTGATTTGGGCATCAAAAGCCTTCCTATACAA
***** ** * **** * * * ***** ***** ** * ** * ** * ** *

10376 TATATTCACCCCAAGAAGCATTAATCAATATCATCCCAAAAAATCTATTCTGTATTAT GACATGTCAAACCTGGGAAATGATCCTAAGATTGCAGAAATCTGTATGTGAAGCTGCTGAG
27567 TATATTCACCCCAAGAGCACTTATCAATATCATCCCAAAAAATCTATTCTGTATTAT GATATGTCCAACTGGGAAATGATCCTCAAAAGTTGCTGAATCAGTCTGTGAAGCTGCAGAG
10292 TATATTCACCCCAAGAAGCATTAATCAATATCATCCCAAAAAATCTATTCTGTATTAT GACATGTCAAACCTGGGAAATGATCCTCAAAAGTCGCTGAATCAGTCTGCGAAGCTGCTCAG
10291 TATATTCACCCCAAGAAGCATTAATCAATATCATCCCAAAAAATCTATTCTGTATTAT GACATGTCAAACCTGGGAAATGATCCTCAAAAGTCGCTGAATCAGTCTGCGAAGCTGCTCAG
***** ***** ** * **** * * * ***** ***** ** * ** * ** * ** *

10376 AGATTTGGATTTTTCAGCTTGTAAACCATGGAGTACTCTTGAAGTTCTTGAGGCTGTT GAGGATGCAACTCATGGTTCTTTGGACTTCGCGCGCCGTAAGAGAACATATTTCTAAA
27567 GAATTTGGCTTCTTTCAGTTGGTTAACCATGGCGTGCCACTTGAGGTACTCGATGGCGTT AAGGATGCAACCCATCGGTTCTTTGGTTGCCAGCAGAGGAGAAGAGAAATTTCCAAA
10292 CAATTTGGCTTCTTTCAGCTGTTTAAACCATGGCGTGCCACTTGAGGTACTCGATGGAGTT AAGGATGCAACCCATCGGTTCTTTGGTTTACCAGCAGAGGAGAAGAGAAATATTTCAAAG
10291 CAATTTGGCTTCTTTCAGCTGTTTAAACCATGGCGTGCCACTTGAGGTACTCGATGGAGTT AAGGACGCAACACATCGTTCTTTGGTTTACCAGCAGAGGAGAAGAGAAATATTTCAAAG
***** ** * **** * * * ***** ***** ** * ** * ** * ** *

10376 GAGCTTTCTCTTCTAACAGTGTCAAGTTTGGCACTAGCTTTAGTCTGATTCTGAGAAA GCTCTTGAATGGAAGACTACCTTAGCCTTTTCTATGTCTCCGACGACGAGCTAATGCA
27567 GAGCTTTCTTCCCAACAACATCAGATTTGGAACCACTTTAGTCTGATGACGAGAGAAA GCTCTTGAATGGAAGATTACCTTAGCCTTTTCTATGTCTCAGAGGAGGAGCTTCTGTA
10292 GAGCTTTCTTCTACCAACAGCGTCAGGTTTGAACCACTTTAGTCTGATGACGAGAGAAA GCTCTTGAATGGAAGGATTATCTCAGCCTCTTTTATGTCTCTGAGGAAGAGCTTTTGCA
10291 GAGCTTTCTTCTACTAACAGCGTCAGGTTTGAACCACTTTAGTCTGATGACGAGAGAAA GCTCTTGAATGGAAGGATTATCTCAGCCTCTTTTATGTCTCTGAGGAAGAGCTTTTGCA
***** ** * **** * * * ***** ***** ** * ** * ** * ** *

10376 TTGTGGCTCTCTCAAGGATGAATGCTTGAATACATGAAGAAAGCAGAAATTTCTA TGCAAAAGCTATTAACTGCCCTAATGGAGAGACTCAATATAAAAGAAATAGATGAAAAG
27567 TTGTGGCTTCTGCTTGCAGAGATGAAGTCTTGAATACATGAAGAAATCTCAAGTCTTC TGACGAAAGCTAATGAGTACACTCATGGAGAACTGAACGTAAAAGAAATAGACGAATCC
10292 TTATGGCTTCTAATGCTTGCAGAGATGAAGTCTTGAATACATGAAGAGATCCCAATCTTT TGCCGAAAGCTCATGAGTCTCTCATGGAGAACTGAACGTGAAGGAAATAGATGAAACA
10291 TTATGGCTTCTAATGCTTGCAGAGATGAAGTCTTGAATACATGAAGAGATCCCAATCTTT TGCCGAAAGCTAATGAGTCTCTCATGGAGAACTGAACGTGAAGGAAATAGATGAAACA
** ***** * **** ***** * ***** ** * ** * ** *

10376 AAAGAGTCTCTTTTAAATGGGATCAAGAAATTAACCTTAATATTATCCAGATGTCCA AACCTCAACTCACCGTCGGGTTGGCCGCCACTCCGACGTCCTCCCTCACTTTTCTC
27567 AAAGAATCTCTGCTAATGGGCTCTAAAAGGATTAATCTCAACTATTATCCAGATGTCCA AACCTCAGCTCACTGTTGGGTTGGCCGCCACTCCGATGTTTCTCTCACTACTTTCTCTC
10292 AAAGAATCTCTTCTAATGGGATCTAAAAGGATCAACCTCAACTACTATCCAGATGTCCG AATCTTAATCTTACTGTCGGGTTGGCCGCCAATCTGATGTTTCTATCTCTCACTTTCTC
10291 AAAGAATCTCTTCTAATGGGATCTAAAAGGATCAACCTCAACTACTATCCAGATGTCCG ** *** ** * ** * ** * ***** ** * ** * ** * ** *

10376 CTCCAAGACGAAATCGCGGGCTTTACGTGCGAGTGAATGAAGGCAAGGGGATGAAGAT GGTGGGTTTCATGTTCTCCCTCCATCAAAAGGATCTCTGTGATTAAATGTTGGCAGTCAATG
27567 CTCAAGACGAAATCGCGGGCTTTACGTGCGAATCAACGAAGCAAGGGGACGACGAAGAT GGCTGGGTTTCATGTTCTCTCTATTGAAGGATCTCTGTGATCAACGTGGGAGATCAATTG
10292 CTCCAAGACGAAATGGTGGACTTTACGTGAGAGTGAATGAAGGCAAGGGGAAAGGAT GGCTGGGTTTCATGTTCTCCCTATTGAAGGATCTCTGTGATAAATGTTGGAGATGCAATTG
10291 CTGCAAGACGAAATCGGTGGACTTTACGTGAGAGTGAATGAAGGCAAGGGGAAAGGAT GGCTGGGTTTCATGTTCTCCCTATTGAAGGATCTCTGTGATTAAATGTTGGAGATGCAATTG
** ***** ** * **** ***** ** * ** * ***** ** * ** * *****
** ***** ** * **** ***** ** * ** * ***** ** * ** * *****

10376 CAGATACTCAGCAATGGTCGATATAGGAGTGTAGAGCACTGTGTGATTGCGAGTGAAGC AAGAATAGAATTTTGATTCCTGTTTTCGTCATCCAAAGCCTACTGATTATCGGACCT
27567 CAGATACTCAGCAATGGTCGATATAGGAGTGTAGAGCACTGTGTGATTGCGAGTGAAGC AAGAACAGAATTTCAATCCCTATTTTCGTAATCCAAAGCCAAATGATGTATCGGTCTC
10292 CAGATACTCAGCAATGGTCGATATAGGAGTGTAGAGCACTGTGTGATTGCGAGTGAAGC AAGAACAGAATTTCAATCCCTATTTTCGTAATCCAAAGCCAAATGATGTATCGGTCCA
10291 CAGATACTCAGCAATGGTCGATATAGGAGTGTAGAGCACTGTGTGATTGCGAGTGAAGC AAGAACAGAATTTCAATCCCTATTTTCGTAATCCAAAGCCAAATGATGTATCGGTCCA
***** ***** ** * **** ***** ** * ***** ** * ***** ** *

10376 TTGCCGGAATGCTTGCCGCCGGAGAGAAGCCTAAATACAAGAATTTTATATTCTGAT TATGTGAAGCATTTCTTCAGGAAGCTCATGACGGAAGAAGACTGTGGAGTTTGTCTGAA
27567 TTGCCTGAGCTGATCGCTGCCGGAGAGAAGCCTAAATATAAGAATCTTTTACTCTGAT TATGTGAAGCATTTCTTCAGGAAGCTCATGATGGCAAGAAGACTGTGGCTTTTGTCTGAG
10292 TTGCCTGAACTGATCGCCGCCGGAGAGAAGCCTAAATATAAAGATATTCTTTACTCTGAT TACGTCAAGCATTTCTTCAGGAAGCTCATGATGGCAAGAAGACTGTGCTTTTGTCTGAG
10291 TTGCCTGAACTGATCGCCGCCGGTGAAGAAGCCTAAATATAAAGACATCTTTTATTCTGAT TATGTCAAGCATTTCTTCAGGAAGCTCATGATGGCAAGAAGACTGTGCTTTTGTCTGAG
**** ** * ** * ** * ***** ***** ** * ** * ***** ** *

10376 GTTGA----- TATGTGAAGCATTTCTTCAGGAAGCTCATGACGGAAGAAGACTGTGGAGTTTGTCTGAA
27567 ATTTCATCGTAA TATGTGAAGCATTTCTTCAGGAAGCTCATGATGGCAAGAAGACTGTGGCTTTTGTCTGAG
10292 ATCTCATCGTAA TACGTCAAGCATTTCTTCAGGAAGCTCATGATGGCAAGAAGACTGTGCTTTTGTCTGAG
10291 ATCTCATCGTAA
* * *

```

**Fig. 8.5 Alignment of the coding DNA sequences of the functioning (expressed) cassava F6'H1 candidate genes.**